

**Identification and Characterization of Genes  
underlying Natural Variation  
in Flowering Time  
in *Arabidopsis thaliana***

**Inaugural-Dissertation**

Zur Erlangung des Doktorgrades  
der Mathematisch-Naturwissenschaftlichen Fakultät  
der Universität zu Köln

vorgelegt von

**Lei Zhang**

aus : Harbin, China

Köln 2017

Berichterstatter: Prof. Dr. George Coupland

Prof. Dr. Ute Höcker

Prüfungsvorsitzender: Dr. Maria Albani

Tag der letzten mündlichen Prüfung: 22 May 2017

## Summary

The time of flowering is crucial for the reproduction success of a plant; it is regulated by a complex network integrating the internal physiological signals and the external environmental signals. *Arabidopsis thaliana* distribute across wide geographic area and experience diverse climatic conditions. To adapt to the local condition, the genes involved in flowering time regulation exhibit great allelic variation in nature, which is often studied using quantitative trait loci (QTL) analysis.

In this thesis, I present the identification and characterization of the genes underlying two flowering time loci identified in a recombinant inbred line (RIL) population derived from a cross between *Ler* and An-1.

### Functional Analysis of the Landsberg *erecta* and Wa-1 alleles of *FRIGIDA*

Most of the natural variation in flowering time in *Arabidopsis* can be attributed to allelic variation at the gene *FRIGIDA* (*FRI*, AT4G00650), which activates expression of the floral repressor *FLOWERING LOCUS C* (*FLC*, AT5G10140). Usually, late-flowering accessions carry functional *FRI* alleles that encode full-length proteins, whereas early flowering accessions contain non-functional alleles that encode truncated version of *FRI* proteins. The two most frequent alleles found in early flowering accessions are the ones present in the common lab strains Columbia (*FRI-Col*) and Landsberg *erecta* (*FRI-Ler*), that are considered null alleles due to a premature stop codon and a deletion of the start codon respectively.

In this chapter, I identified *FRI* as the causal gene of the flowering QTL on the top of chromosome 4 in the *Ler* x An-1 RIL population. I further analyzed flowering time data from various sources and found evidences that *FRI-Ler* retains some functionality with respect to the truly null *FRI-Col* allele, challenging the current functional classification for *FRI*. To test this hypothesis, I functionally tested a number of *FRI* alleles from different *Arabidopsis* accessions by cloning and transforming them in a common genetic background and characterizing the molecular and physiological variation induced in the resulting lines. In contrast to common belief, *FRI-Ler* is able to upregulate the expression of *FLC* and delay flowering time; though its effect is weaker

compared to the fully functional *FRI* alleles due to its reduced expression level. In addition, I identified a novel non-functional allele present in the accession Wa-1 that encodes a full-length FRI protein whose stability is affected by a single amino acid mutation.

In summary, the findings presented in this chapter increase the accuracy of the current functional classification for *FRI*, one of the largest contributors to the natural variation in flowering time found in *Arabidopsis*.

### **Identification of *VIP HOMOLOG2 (VIH2)* as a novel flowering time gene underlying a flowering time QTL**

This chapter focuses on the flowering time locus at the top of chromosome 3 (named QTL3).

I studied the presence of QTL3 in various segregating populations and found it in three RIL populations that had *Ler* as a common parent (*Ler* x An-1, *Ler* x Sha and *Ler* x Eri). QTL3 was confirmed in the Near Isogenic Lines (NILs) derived from all three; the *Ler* allele at QTL3 delayed flowering in all cases. Fine mapping of QTL3 with the *Ler* x Eri NILs narrowed down the locus to a single gene: *VIP HOMOLOG 2 (VIH2)*, AT3G01310). *VIH2* encodes a kinase that catalyzes the synthesis of inositol pyrophosphates (InsP<sub>7</sub> and InsP<sub>8</sub>) and is well conserved in all eukaryotic organisms. We measured InsP<sub>8</sub> in the NILs and found higher abundance associated to the *VIH2-Ler* allele. Mutant lines for *vih2* are early flowering, further proving the role of *VIH2* in flowering time regulation, and implicating that *VIH2-Ler* is a gain-of-function mutant allele. In agreement with this, we found that the *VIH2-Ler*-NIL is also more responsive to jasmonate signaling, a pathway that linked InsP<sub>8</sub> signaling to flowering time.

# Zusammenfassung

Die Zeit der Blüte ist entscheidend für den Reproduktionserfolg einer Pflanze; der Beginn der Blüte wird durch ein komplexes Netzwerk geregelt, das die internen physiologischen Signale und die äußeren Umgebungssignale integriert. *Arabidopsis thaliana* ist über ein weites geografisches Gebiet verbreitet und erlebt vielfältige klimatische Bedingungen. Um sich an die örtliche Bedingungen anzupassen, zeigen die Gene, die an der Blütezeitregulation beteiligt sind, eine große Allelvariation in der Natur, die häufig mit Analysen der Region eines quantitativen Merkmals (QTL) untersucht wird.

Ich reanalysierte die Blütezeit QTL-Analyse, die in einer rekombinanten Inzuchtlinie (RIL) - Population durchgeführt wurde, die von einer Kreuzung zwischen *Ler* und An-1 abgeleitet wurde. In dieser Arbeit präsentiere ich die Identifizierung und Charakterisierung der Gene, die zwei QTLs in dieser Studie zugrunde liegen in zwei separaten Kapiteln.

## Funktionsanalyse der Landsberg *erecta* und Wa-1 Allele von FRIGIDA

Die meisten der natürlichen Variation der Blütezeit in *Arabidopsis* lassen sich auf die Allelvariationen im Gen *FRIGIDA* (*FRI*, AT4G00650) zurückführen, das die Expression des Blütenrepressors *FLOWERING LOCUS C* (*FLC*, AT5G10140) aktiviert. Gewöhnlich tragen spätblühende Akzessionen funktionale *FRI*-Allele, die Volllängen-Proteine codieren, während früh blühende Akzessionen nicht-funktionelle Allele enthalten, die eine gekürzte Version von *FRI*-Proteinen kodieren. Die beiden häufigsten Allele, die in den früh blühenden Akzessionen gefunden wurden, sind die, die in den üblichen Laborstämmen Kolumbien (*FRI-Col*) und Landsberg *erecta* (*FRI-Ler*) vorhanden sind, welche aufgrund eines vorzeitigen Stopcodons und eines fehlenden Startcodons als Nullallele betrachtet werden.

In diesem Kapitel habe ich *FRI* als das dem Blütezeit QTL am Anfang von Chromosom 4 in der *Ler* x An-1 RIL-Population zugrundeliegende Gen identifiziert. Ich analysierte Blütezeit-Daten aus verschiedenen Quellen und stellte fest, dass *FRI-Ler* einige Restfunktionalität gegenüber dem wirklichen Nullallel *FRI-Col*- behält, und damit die derzeitige funktionale Klassifizierung für *FRI* infrage stellt. Um diese Hypothese zu testen, habe ich funktionell eine Anzahl von *FRI*-Allelen aus verschiedenen *Arabidopsis*-Akkzessionen durch Klonierung und Transformation in einem

gemeinsamen genetischen Hintergrund getestet und die molekulare und physiologische Variation, die in den resultierenden Linien induziert wird, charakterisiert. Im Gegensatz zur derzeitigen Annahme ist *FRI-Ler* in der Lage, die Expression von *FLC* zu erhöhen und die Blütezeit zu verzögern - wobei die Wirkung im Vergleich zu den voll funktionsfähigen *FRI*-Allelen aufgrund des reduzierten Expressionsniveaus schwächer ist. Darüber hinaus habe ich ein neuartiges nicht-funktionelles Allel identifiziert, das in der Akzession Wa-1 vorhanden ist, welches ein FRI-Protein in voller Länge kodiert, dessen Stabilität durch eine einzige Aminosäuremutation beeinflusst wird.

Zusammengenommen verbessern die in diesem Kapitel präsentierten Ergebnisse die Genauigkeit der aktuellen funktionalen Klassifikation für FRI, einem der Hauptakteure der natürlichen Variation der Blütezeit in Arabidopsis.

### **Identifizierung von *VIP HOMOLOG 2 (VIH2)* als neuartiges Blütezeit-Gen, das einem Blütezeit QTL zugrunde liegt**

Dieses Kapitel konzentriert sich auf die Region eines Blütezeitmerkmals am Anfang von Chromosom 3 (bezeichnet als QTL3).

Ich untersuchte die Anwesenheit von QTL3 in verschiedenen segregierenden Populationen und fand sie in drei RIL-Populationen, die Ler als gemeinsames Elternteil (*Ler* x An-1, *Ler* x Sha und *Ler* x Eri) hatten. QTL3 wurde in den von allen drei abgeleiteten nahezu isogenen Linien (NILs) bestätigt; Das *Ler*-Allel bei QTL3 verzögert die Blüte in allen Fällen. Die Feinkartierung von QTL3 mit den *Ler* x Eri NILs grenzt den Locus auf ein einziges Gen ein: *VIP HOMOLOG 2 (VIH2)*, AT3G01310). *VIH2* kodiert eine Kinase, die die Synthese von Inositolpyrophosphaten (InsP7 und InsP8) katalysiert und in allen eukaryotischen Organismen gut konserviert ist. Wir bestimmten InsP8 in den NILs und fanden eine erhöhte Menge, die mit dem *VIH2-Ler*-Allel assoziiert war. Mutante Linien für *vih2* sind früh blühend, was weiterhin die Rolle von *VIH2* in der Blütezeitregulation belegt, und impliziert, dass *VIH2-Ler* eine Funktionsgewinn-Mutante ist. In Übereinstimmung damit haben wir festgestellt, dass die *VIH2-Ler-NIL* auch auf Jasmonat reagiert, ein Signalweg, der die InsP8-Signalisierung mit der Blütezeit verknüpft.

## Table of contents

### Chapter 1

<b>General introduction on flowering time regulation in <i>Arabidopsis thaliana</i> and quantitative trait loci analysis</b> .....	1
<b>1.1 Regulation of flowering time in <i>Arabidopsis</i></b> .....	2
1.1.1 Vernalization and autonomous pathway .....	2
1.1.2 Photoperiodic pathway .....	4
1.1.3 Thermosensory pathway .....	6
1.1.4 Age pathway .....	7
1.1.5 Hormone pathway .....	9
<b>1.2 Quantitative trait loci analysis (QTL) and its application on natural variation studies</b> .....	11
1.2.1 QTL analysis .....	11
1.2.2 Recombinant inbred line (RIL) as mapping population .....	11
1.2.3 Fine mapping and confirmation of candidate genes underlying a QTL .....	12

### Chapter 2

<b>Functional Analysis of the Landsberg <i>erecta</i> and Wa-1 alleles of <i>FRIGIDA</i></b> .....	14
<b>2.1 INTRODUCTION</b> .....	15
<b>2.2 RESULTS</b> .....	18
2.2.1 Analysis of published data suggest <i>FRI-Ler</i> functionality .....	18
2.2.2 <i>FRI-Ler</i> may encode a N-terminus truncated functional protein .....	19
2.2.3 Functionality of <i>FRI-Ler</i> is confirmed in transgenic plants .....	21
2.2.4 The <i>FRI-Ler</i> allele presents expression defects due to cis-regulation .....	23
2.2.5 Lack of effect of coding polymorphisms in <i>FRI-Ler</i> .....	25
2.2.6 Fitness effect of the different <i>FRI</i> alleles .....	26
2.2.7 A new non-functional <i>FRI</i> allele: <i>FRI-Wa-1</i> .....	28
2.2.8 Single amino acid mutation abolishes the function of <i>FRI-Wa-1</i> allele .....	31
2.2.9 Analysis of other accessions with <i>FRI-Wa-1</i> allele .....	34
<b>2.3 DISCUSSION</b> .....	35

<b>2.4</b>	<b>MATERIAL AND METHODS</b>	38
2.4.1	Analysis of published datasets	38
2.4.2	Sequencing of <i>FRI</i> alleles	39
2.4.3	Cloning of <i>FRI</i> alleles	39
2.4.4	Selection and phenotyping of transgenic plants	41
2.4.5	Expression analysis using quantitative real-time PCR	43
2.4.6	Allele-specific expression analysis	43
2.4.7	Microscopic imaging	44
2.4.8	Western blot assay	44

## Chapter 3

### Identification of *VIP HOMOLOG2* (*VIH2*) as a novel flowering time gene

	underlying a flowering time QTL	46
<b>3.1</b>	<b>INTRODUCTION</b>	47
<b>3.2</b>	<b>RESULTS</b>	49
3.2.1	Confirmation of QTL3 in multiple populations	49
3.2.2	Characterization of the <i>Ler</i> x <i>Eri</i> NILs	52
3.2.3	Identification of <i>VIH2</i> as causal gene underlying QTL3	54
3.2.4	Protein sequence and functional variation between <i>VIH2-Eri</i> and <i>VIH2-Ler</i> alleles	57
3.2.5	<i>VIH2-Ler</i> NIL is more sensitive to jasmonate	59
3.2.6	<i>VIH2</i> allele variation and distribution among Arabidopsis accessions	61
<b>3.3</b>	<b>DISCUSSION</b>	63
3.3.1	Identification of <i>VIH2</i> as the causal gene of flowering time QTL3	63
3.3.2	Function of <i>VIH2</i> and its product: inositol pyrophosphates in plants	64
3.3.3	The role of <i>VIH2</i> in flowering time regulation	65
3.3.4	Natural variation of <i>VIH2</i> among Arabidopsis accessions	66
3.3.5	Pleiotropic effect of inositol pyrophosphates in yeast and animals	68
3.3.6	Outlook for pleiotropy of <i>VIH2</i> in Arabidopsis	69
<b>3.4</b>	<b>Materials and Methods</b>	71
3.4.1	Plant materials and phenotyping of flowering time	71
3.4.2	Genotyping at QTL3	71



3.4.3	Expression analysis using RNA-seq and quantitative real-time PCR.....	71
3.4.4	HPLC Analyses of Inositol Phosphates .....	72
3.4.5	Root Length Measurement .....	72
3.4.6	Analysis on <i>VIH2</i> allele variation and distribution .....	72
3.4.7	Accession Numbers.....	73
<b>REFERENCES.....</b>		<b>76</b>
<b>Curriculum vitae.....</b>		<b>87</b>
<b>Acknowledgements .....</b>		<b>88</b>
<b>Erklärung.....</b>		<b>89</b>

## List of Figures

<b>Figure 2-1.</b>	
QTL study on flowering time of a RIL population derived from a cross between An-1 and Ler accessions (Tisne et al., 2010).....	17
<b>Figure 2-2.</b>	
Evidence of FRI-Ler functionality in published data.....	19
<b>Figure 2-3.</b>	
Genomic and protein sequence comparison between FRI-Ler and FRI-Col to FRI-Bil-7. ....	20
<b>Figure 2-4.</b>	
Characterization of transgenic lines carrying the FRI-Col or FRI-Ler allele. ....	22
<b>Figure 2-5.</b>	
Expression of FRI-Ler is lower in Arabidopsis accessions and in F <sub>1</sub> hybrids. ....	24
<b>Figure 2-6.</b>	
Flowering time of transgenic lines carrying promoter-swap constructs involving the FRI-Ler and FRI-Bil-7 alleles. ....	26
<b>Figure 2-7.</b>	
Fitness related traits of accessions carrying different FRI alleles growing in field. ....	27
<b>Figure 2-8.</b>	
Flowering time of transgenic plants carrying selected alleles of FRI. ....	29
<b>Figure 2-9.</b>	
Characterization of transgenic lines carrying the FRI-Wa-1 or FRI-Bil-7 alleles. ....	31
<b>Figure 2-10.</b>	
Characterization of L294F amino acid mutation in FRI protein. ....	33
<b>Figure 2-11.</b>	
Collection sites of accessions carrying same FRI allele as FRI-Wa-1.....	34
<b>Figure 2-12.</b>	
Selection of T <sub>1</sub> transgenic plants.....	42
<b>Figure 3-1.</b>	
Confirmation of QTL3 in three NILs.....	50
<b>Figure 3-2.</b>	
Interaction between QTL3 and top of chromosome 5 in the Ler x An-1 RIL population. ....	51

<b>Figure 3-3.</b>	
RNA-seq analysis of Eri-NIL and Ler-NIL. ....	53
<b>Figure 3-4.</b>	
Effect of vernalization on the flowering time of Eri and Ler-NILs. ....	54
<b>Figure 3-5.</b>	
Identification of VIH2 as causal gene underlying QTL3 .....	55
<b>Figure 3-6.</b>	
VIH2 delays flowering in the Col background. ....	57
<b>Figure 3-7.</b>	
Sequence alignment of VIP homolog proteins.....	58
<b>Figure 3-8.</b>	
Ler encodes an enzymatically more active allele of VIH2. ....	59
<b>Figure 3-9.</b>	
Root growth of VIH2-NILs on MeJA containing agar medium. ....	60
<b>Figure 3-10.</b>	
Geographic distribution of the common VIH2 alleles. ....	62
<b>Figure 3-11.</b>	
Proposed molecular mechanism of VIH2 regulation on flowering. ....	66
<b>Figure 3-12.</b>	
Synthesis of inositol pyrophosphate from InsP <sub>6</sub> in yeast and animal. ....	69

## List of Tables

<b>Table 2-1.</b>	
The species and GenBank ID of FRI protein sequences used in the alignment. ....	38
<b>Table 2-2.</b>	
Position and sequence of the primers used for cloning. ....	39
<b>Table 2-3.</b>	
Position and sequence of the primers used for cloning. ....	41
<b>Table 2-4.</b>	
Primers used in the real-time PCR.....	43
<b>Table 2-5.</b>	
Primers used in the allele specific analysis.....	44
<b>Table 3-1.</b>	
Primers used in for qRT-PCR.....	72
<b>Table 3-2.</b>	
The species and GenBank ID of VIP homologs protein sequences used in the alignment. ....	73
<b>Table 3-3.</b>	
Primers used for fine mapping of QTL3.....	74
<b>Table 3-4.</b>	
Primers used in Sanger sequencing of <i>VIH2</i> locus. Position of the primer is relative to the start codon of <i>VIH2</i> .....	75

## Chapter 1

---

# General introduction on flowering time regulation in *Arabidopsis thaliana* and quantitative trait loci analysis

### Summary

It is essential for plants to coincide their flowering with favorable conditions for production of fruits and seeds to enhance their reproductive success. Flowering at the right time requires responding to a diverse range of environmental and internal signals and integrating the signals into a single decision: floral transition, the switch from vegetative to reproductive phase.

The regulation of flowering time is most extensively studied in the model organism *Arabidopsis thaliana*. The identified flowering time genes are integrated into six distinct but interlinked genetic pathways, known as vernalization, autonomous, photoperiod, temperature, age and hormone pathways. In addition, studies have shown links between flowering time and other factors, including sugar metabolism, nutrient availability, drought and herbivore predation. All the genetic pathways form a complex network that enable the plant to monitor its external and internal conditions, and converge the information to regulate floral integrators such as *FLOWERING LOCUS T (FT)*, *LEAFY (LFY)* and *SUPPRESSOR OF OVEREXPRESSION OF CO 1 (SOC1)*, which promote the decision to floral transition.

*Arabidopsis* distribute across a wide range of latitude and altitude; thus face diverse environmental conditions. The regulation on flowering time reflects adaptation to the populations' local environment; consequently, diverse natural variations exist in flowering time regulators. Quantitative trait loci (QTL) analysis is a tool frequently used to identify the genetic basis of such variation. This chapter briefly reviews the current knowledge on the flowering time regulation and the principle of QTL analysis, which is often exploited to study the natural variation in flowering time genes.

## 1.1 Regulation of flowering time in Arabidopsis

### 1.1.1 Vernalization and autonomous pathway

Many plants actively repress flowering until they experience a prolonged cold period (vernalization); this prevents flowering and seed production during the unfavorable winter condition. The main player in the vernalization pathway is a major flowering repressor *FLOWERING LOCUS C (FLC)*. *FLC* encodes a MADS-box transcription factor, which complexes with *SHORT VEGETATIVE PHASE (SVP)* to form a potent transcriptional repressor of floral inducers, such as *FLOWERING LOCUS T (FT)* and *SUPPRESSOR OF CONSTANS 1 (SOC1)* (Amasino, 2010; Michaels and Amasino, 1999). Short term (2 to 3 weeks) exposure to cold temperature induces the expression of *COOLAIR*, a set of antisense non-coding transcripts near the poly-A site of the *FLC* coding region and causes down-regulation of *FLC* (Song et al., 2012a; Swiezewski et al., 2009). Longer exposure to cold temperatures induces the transcription of *COLD AIR*, a non-coding sense transcript from the first intron, that guides the histone methyltransferase subunit of the Polycomb Repressive Complex2 (PRC2) to the *FLC* locus (Heo and Sung, 2011). The quantitative accumulation of the Polycomb-based complexes and histone modifications during vernalization leads to a stable epigenetic silencing of the *FLC* locus that is maintained upon plants' return to warm conditions (Angel et al., 2011; De Lucia et al., 2008). The expression level of *FLC* is re-set during embryogenesis via chromatin modification; ensuring that every generation of newly germinated seedlings require vernalization to flower despite coming from previously vernalized parents (Choi et al., 2009; Sheldon et al., 2008).

*FRIGIDA (FRI)* is the major inducer of *FLC*, acting as a scaffold to form a transcription activation complex by interaction with *FRIGIDA LIKE1 (FRL1)* at the N-terminus and *SUPPRESSOR OF FRIGIDA4 (SUF4)*, *FRIGIDA ESSENTIAL1 (FES1)* and *FLC EXPRESSOR (FLX)* at the C-terminus (Choi et al., 2011). This complex interacts with conserved chromatin modification factors such as the SWR1 chromatin remodeling complex (SWR1-C) and *EARLY FLOWERING IN SHORT DAYS (EFS)* (Ko et al., 2010) to increase the level of H3K4me3, H3K36me3 and acetylation on H3 and H4 at the *FLC* chromatin region; therefore leading to high transcription of *FLC* (Ding et al., 2013). Furthermore, *FRI* is suggested to be involved in co-transcriptional processes which link the function of 5' end capping with transcription and efficient

splicing of *FLC* (Geraldo et al., 2009). The *FRI* transcript can be first detected during embryogenesis, and continues to be present in all tissues throughout the lifetime of Arabidopsis. This is also true for the FRI protein, although this protein is degraded during cold treatment; indicating that FRI also plays a role in down-regulating *FLC* during vernalization (Hu et al., 2014). Extensive natural variation exists in *FRI* among natural populations; the *FRI* alleles can be categorized according to their function. Plants with functional *FRI* allele usually have elevated expression of *FLC* and require vernalization to flower; in contrast, plants carrying non-functional allele of *FRI* have low level of *FLC* and are early flowering without vernalization (Johanson et al., 2000).

Autonomous pathway can repress the expression of *FLC* independent of vernalization through RNA-processing mechanisms and/or by chromatin modification (Simpson, 2004). *FLOWERING LOCUS CA* (*FCA*) and *FLOWERING LOCUS PA* (*FPA*) contain RNA-recognition motifs (RRM) proteins; they function partially redundantly to control alternative splicing and 3'-end processing *COOLAIR* (Hornyik et al., 2010; Liu et al., 2007). *COOLAIR* is a collection of long noncoding antisense transcripts that cover the *FLC* locus; its promoter is downstream of the polyadenylation (poly(A)) site of the *FLC* sense transcript and its transcripts can terminate either at proximal or distal poly(A) sites. Proteins of FCA and FPA directly associate with *FLC* chromatin near the *COOLAIR* proximal poly(A) site; they act together with *FLOWERING LOCUS Y* (Abulencia et al.), a RNA 3'-end processing factor, to promote the choice of the proximal poly(A) site in *COOLAIR* (Liu et al., 2010). The elevated ratio between proximal to distal *COOLAIR* transcripts triggers the activity of histone demethylase *FLOWERING LOCUS D* (*FLD*) (Liu et al., 2007). *FLD* forms a chromatin modification complex with *FLOWERING LOCUS VE* (*FVE*) and *HISTONE DEACETYLASE 6* (*HDA6*); together, they suppress the level of H3K9K14Ac and H3K4Me3 at the *FLC* locus and repress the expression of *FLC* (Yu et al., 2016). The autonomous pathway also comprises the homeodomain protein *LUMINIDEPENDENS* (*LD*) (Lee et al., 1994) and the K homology-domain protein *FLOWERING LATE WITH KH MOTIFS* (*FLK*) (Lee et al., 1994; Lim et al., 2004), they both repress the expression of *FLC* likely through RNA process, but how their roles fit in the above network remain to be studied.

The repression of *FLC* by the autonomous pathway can be over-ridden by the function of *FRI*, which up-regulates *FLC*. However, in the null-*fri* genetic background such as Col, mutation in any autonomous pathway genes causes elevated *FLC* level and delayed flowering, mimicking the phenotype of plants with functional *FRI*. Similarly, the late flowering phenotype is eliminated by mutation of *FLC* (Michaels and Amasino, 2001) or vernalization. In summary, both vernalization and autonomous pathways converge on *FLC*, which suppresses downstream floral integrators to regulate flowering time.

### 1.1.2 Photoperiodic pathway

*Arabidopsis* distribute mainly in temperate zone, where changes in day length occur in a predictable seasonal manner. Photoperiodic pathway ensures *Arabidopsis* flower earlier in the summer long days (LD) than in the winter short days (SD) (Kobayashi and Weigel, 2007). *CONSTANS* (*CO*), a major transcription activator of *FT* and *SOC1*, is at the core of the photoperiodic pathway. Complex information from the intrinsic circadian clock and the external light environment is integrated to regulate *CO* at both transcriptional and post-translational level (Suarez-Lopez et al., 2001; Valverde et al., 2004).

The mRNA abundance of *CO* oscillates throughout the day under the control of the intrinsic circadian clock. In the morning, clock genes *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*) and *LATE ELONGATED HYPOCOTYL* (*LHY*) induce the expression of the *CYCLING DOF FACTORS* (*CDFs*) (Schaffer et al., 1998; Wang and Tobin, 1998). The CDF proteins bind to the *CO* locus and suppress its transcription (Fornara et al., 2009). In the afternoon, the transcription of *CDFs* is inhibited by the *PSEUDO-RESPONSE REGULATOR* (PRR) proteins (Nakamichi et al., 2007). In addition, the CDF proteins are degraded by a complex of FLAVIN-BINDING KELCH REPEAT F-BOX 1 (FKF1) and GIGANTEA (GI). The formation of FKF1-GI complex is blue-light dependent and its abundance depends on the photoperiod. In LD, the peaks of FKF1 and GI coincide in the late afternoon; allowing accumulation of the complex, which degrade CDF proteins and release *CO* from transcriptional repression (Sawa et al., 2007). In SD, GI peaks at dusk but FKF1 peaks after dark, so the abundance of FKF1-GI complex is significantly reduced. Consequently, the transcriptional repression of *CO* and *FT* cannot be efficiently removed. In summary, the *CO* expression rises around 12 h after dawn and stays high until the following dawn independent of



photoperiod (Suarez-Lopez et al., 2001); but in LD condition, *CO* expression is further promoted by the activity of FKF1-GI complex.

*CO* transcripts accumulate at night; however, *CO* protein level does not follow the same pattern due to light dependent post-translational regulation. In general, the *CO* protein is stabilized in blue and far-red light but destabilized in red light or in the dark (Valverde et al., 2004). The RING-finger E3 ubiquitin ligase CONSTITUTIVE PHOTOMORPHOGENIC 1 (COP1) is the major negative regulator of *CO* protein; it forms a complex with SUPPRESSOR OF PHYA-105 (SPA1) family members and degrades *CO* protein in the dark. At daytime, multiple photoreceptors involve in *CO* stability regulation. Cryptochrome 1 (CRY1) and 2 (CRY2) inhibit the COP1-SPA complex in the presence of blue light (Laubinger et al., 2006; Liu et al., 2008b). Phytochrome A (PHYA), the far-red light photoreceptor, stabilizes *CO* protein under a low red to far-red (R:FR) ratio, around dawn and dusk (Kim et al., 2008). The red light photoreceptor PHYB recruits the E3 ubiquitin ligase HIGH EXPRESSION OF OSMOTICALLY RESPONSIVE GENES 1 (HOS1) to degrade *CO* protein at the high R:FR ratio during the day (Lazaro et al., 2012). Circadian genes are also involved in the post-translational regulation of *CO*, they contribute to the recognition of LD and allow the *CO* accumulation specifically in LD condition. GI binds to ZEITLUPE (ZTL) in the afternoon and stops the ZTL mediated *CO* degradation (Kim et al., 2007; Song et al., 2014). Meanwhile, the FKF1-GI complex, which only peaks at dawn in the LD, stabilizes the *CO* protein (Song et al., 2012b). PPR proteins also stabilize *CO* protein by direct interaction, specifically in the morning and before dawn at LD condition (Hayama et al., 2017).

The overall result of the multi-layered regulation on *CO* is the photoperiod dependent oscillation of protein accumulation. Under SD condition, the *CO* protein level is maintained under low level, whereas under LD condition, *CO* protein level peaks briefly after dusk and degrades during the day until it peaks a second time from mid-day to dawn and then degrades in the dark (Valverde et al., 2004). Furthermore, CDF1 associates with the *FT* promoter at the *CO* binding site, to repress the expression of *FT* in the morning. Similar to the transcriptional regulation of *CO*, the FKF1-GI complex degrades CDF1 in the afternoon and enables the activation of *FT* by *CO* (Song et al., 2012b). Therefore, high *FT* transcription coincides with the *CO* peak in the afternoon under LD condition, whereas very low levels of *FT* transcripts can be found in SD condition. *FT*

transcripts accumulate in the companion cells of the leaves. They encode mobile proteins that are transported through the phloem to the shoot apex (Corbesier et al., 2007). Upon its arrival at the shoot apical meristem, the FT protein activates the expression of downstream genes such as *SOC1*, *FRUITFULL (FUL)*, *APETALA 1 (API)* and *LEAFY (LFY)*; and leads to the formation of an inflorescence meristem and to the production of flowers (Torti et al., 2012; Wigge et al., 2005; Yoo et al., 2005).

### 1.1.3 Thermosensory pathway

The ambient temperature is another important environmental factor for *Arabidopsis* to regulate flowering. The long-term exposure to low temperature is studied in the vernalization pathway; the thermosensory pathway mainly studies the flowering inducing effect of warm growth temperature: an elevation of growth temperature from 22 °C, which is the common greenhouse temperature, to 27 °C can induce the flowering of *Arabidopsis* plants despite the non-inductive SD photoperiod (Balasubramanian et al., 2006). Research has shown that the floral induction by high ambient temperature is achieved through several independent mechanisms, as reviewed by (McClung et al., 2016).

*Arabidopsis* employs the histone variant H2A.Z as a thermosensor. Less H2A.Z is incorporated into nucleosomes at higher temperature, resulting in better accessibility of transcription factors on the DNA and higher gene expression (Talbert and Henikoff, 2014). The *FT* locus is subject to the regulation of H2A.Z; at elevated temperature, H2A.Z is depleted from the *FT* promoter region and allows the binding of PHYTOCHROME INTERACTING FACTOR proteins (PIF4 and PIF5) to up-regulate the expression of FT (Fernandez et al., 2016; Kumar et al., 2012). PIF4 physically interacts with CO protein, but their effect on *FT* expression upon temperature increase are additive (Fernandez et al., 2016). Plants with loss-of-function mutation in *ACTIN-RELATED PROTEIN6 (ARP6)* are defective in H2A.Z deposition to chromosomes; however, they can still accelerate their flowering in response to increased temperature (Galvao et al., 2015), indicating mechanisms independent of H2A.Z.

*SVP*, the flowering repressor involved in vernalization pathway, plays a crucial role in the thermosensory pathway. *SVP* forms transcriptional repressing complex with *FLC*, as well as related MADS-box transcription factors *FLOWERING LOCUS M (FLM)* and *MADS AFFECTING*

*FLOWERING 2 (MAF2)*. The transcripts from both *FLM* and *MAF2* are subjected to temperature dependent alternative splicing. At elevated temperature, the splice variant of both genes are unable to form a functional complex with *SVP*, thus fail to inhibit the expression of *FT* - resulting in earlier flowering (Airoldi et al., 2015; Lee et al., 2013; Pose et al., 2013). Moreover, the fluctuating daily temperature ranging between 12 to 32 °C is inhibitory to the expression of *FLC*, which explains why the effect of *FRI*-mediated late flowering phenotype is less prominent in the field condition than in the stable greenhouse condition (Burghardt et al., 2016). In addition, loss of function mutants of the autonomous pathway genes *FVE*, *FCA* and *FLD*; together with a gain-of-function mutant of *CRY2* do not show early flowering in response to elevated temperature (Balasubramanian et al., 2006). Involvements of these genes in thermosensory pathway implicate the complex cross talks between thermosensory pathway and autonomous pathway and vernalization pathway, in addition to the well-studied link to the photoperiod pathway.

#### 1.1.4 Age pathway

The external environment is crucial for plants to determine their flowering time. However, a plant in its juvenile phase cannot flower even in the inductive environment; whereas plants old enough may flower eventually even in the non-inductive environments. This age-dependent flowering regulation, the age pathway, is orchestrated by two groups of microRNAs (miRNA), miR156/157 and miR172 (as reviewed by Teotia and Tang (2015)).

The miR156 is necessary and sufficient for maintaining the juvenile phase in plants. The expression level of miR156 is highest in the young seedlings and gradually decreases as the plants grow and develop toward flowering (Wang et al., 2009; Wu and Poethig, 2006). Sugar serves as an endogenous cue promoting the juvenile-to-adult phase transition (Roldan et al., 1999). It was found recently that sucrose, glucose, or maltose down-regulates the expression of primary transcripts of all miR156 members studied, except miR156B and miR156D (Yang et al., 2013; Yu et al., 2013). The sugar-dependent regulation of miR156 is maybe facilitated by the glucose based signaling molecule, rehalose-6-phosphate (T6P). The accumulation of T6P positively correlates with the sucrose level and negatively regulates the expression of miR156 (Wahl et al., 2013). The miRNA regulates its target genes by suppressing their expression through both post-transcriptional gene silencing and translational inhibition. The targets of miR156 are members of the transcription factor

protein family named *SQUAMOSA PROMOTER-BINDING PROTEIN-LIKEs* (*SBPs/SPLs*). Their expression increases as the expression of miR156 declines, thus promoting the transition from juvenile phase to adult phase (Gandikota et al., 2007; Wang et al., 2009; Wu and Poethig, 2006). The SPL proteins function redundantly as flowering promoters; but based on the size and functions, they can be divided into two groups: SPL3/4/5 encode small proteins and function mainly in flowering time regulation; SPL9/15 encode larger proteins and involve in leaf initiation and flowering regulation (Yu et al., 2013). Recently it was found that *SPL15* inhibits the expression of miR156B, whereas *SPL9/10* up-regulate the expression of miR156; the findings suggest that miR156 and SPL proteins form a complex feedback and feed forward loop in order to regulate the phase transition from juvenile to flowering adult (Wei et al., 2012; Wu et al., 2009).

The other major player in the age pathway is miR172; it acts downstream of miR156 in an antagonistic expression pattern. The expression of miR172 is up regulated by *SPL9* and *SPL10*, so it increases as the plant grow from juvenile stage to flowering stage. miR172 inhibits the expression of the *APETALA 2* (*AP2*) transcription factor family, including *AP2*, *TARGET OF EAT1* (*TOE1*), *TOE2*, *TOE3*, *SCHLAFMUTZE* (*SMZ*), and *SCHNARCHZAPFEN* (*SNZ*); all flowering repressors (Aukerman and Sakai, 2003; Chen, 2004; Schwab et al., 2005). The *AP2* transcription factors also form a feedback loop with miR172; *AP2* inhibit the expression of miR172 whereas *TOE1* and *TOE2* promote the expression of miR172.

Although the decline in expression of miR156, the central player in the age pathway, is only dependent on the age of plant, the downstream members in the age pathway can be regulated by other signals. At juvenile stage, the expression of *SPL3/4/5* are higher in plants in LD condition than in SD condition, though at later stages of vegetative development, expression of *SPL3/4/5* is induced irrespective of photoperiod (Jung et al., 2012; Schmid et al., 2003). In addition, under SD condition, expression of *SPL* genes is up regulated in response to the plant hormone gibberellic acid (GA) (Yu et al., 2012). At lower ambient temperature (16 °C), the *SPL* genes are more sensitive to the inhibition of miR156 (Kim et al., 2012). Furthermore, the miR172 and its target *AP2* family genes interact with members from other flowering time pathways (Cho et al., 2012; Galvao et al., 2012; Jung et al., 2012). The overall effect is that the age pathway safeguards the flowering of adult plants, regardless of environmental factors. However the crosstalk between the age pathway and

other environment-dependent flowering pathways enable that flowering occurs at favorable external conditions.

### 1.1.5 Hormone pathway

Plant hormones (phytohormones) are endogenously occurring compounds; they are present in the plants at very low concentrations, but regulate various aspects of plants' growth and development (Davies, 2004). Major phytohormones include cytokinins, auxins, gibberellic acid (GA), abscisic acid (ABA), brassinosteroids (BRs), jasmonic acid (JA) and ethylene; many of them have been implicated to be involved in flowering time regulation (Davis, 2009).

GA is the best-studied phytohormone in terms of flowering time regulation. Arabidopsis mutants with defects in GA signaling pathway show delayed flowering, especially in the non-inductive SD condition. In contrast, mutants with enhanced GA-signaling and plants subject to exogenous addition of GA are early flowering (Koornneef et al., 1991). Upon binding of GA molecules to its receptors *GA INSENSITIVE DWARF1* (*GID1*), the *GID1* proteins undergo conformational change, which leads to the binding and degradation of DELLA proteins. The effect of GA on flowering is mainly through the five DELLA proteins, which function redundantly as flowering repressors in Arabidopsis (Griffiths et al., 2006; Murase et al., 2008; Willige et al., 2007). Flowering promoter genes *SOC1* and *AGL24* respond to GA treatment with increased expression, yet flowering repressor *SVP* showed decreased expression (Li et al., 2008; Liu et al., 2008a). GA is also known to regulate the flowering time integrator *LFY*, through the miR159 mediated regulation of *GAMYB* transcription factors (Achard et al., 2004).

The role of ABA as a flowering repressor was implicated by the early flowering phenotype of the ABA deficient mutants (Chandler et al., 2000). Recent studies provide a mechanistic explanation: *ABSCISIC ACID-INSENSITIVE 4* (*ABI4*), a transcription factor involved in ABA signaling pathway, up-regulate the expression of flowering repressor *FLC* by binding to the promoter region of *FLC* gene upon activation by ABA (Shu et al., 2016). In contrast to ABA, BRs are proposed to be flowering promoters because of the late-flowering phenotype of the mutants deficient in BR biosynthesis (Azpiroz et al., 1998; Chory et al., 1991). Further evidences come from the early flowering mutant unable to inactivate BRs and the late flowering mutant deficient in BR receptor (Domagalska et al., 2007; Turk et al., 2005). It's been shown that BRs promote

flowering by suppressing the expression of *FLC* via the autonomous pathway (Domagalska et al., 2007; Yu et al., 2008). Jasmonate is another phytohormone whose role in flowering time regulation has been studied. Sheard et al. (2010) showed that the JASMONATE-ZIM DOMAIN (JAZ) proteins inhibit the flowering repressors *TOE1* and *TOE2*, preventing them from inhibiting *FT* transcription. Upon activation of JA pathway, the JAZ proteins are degraded and the *TOE1/2* proteins are free to inhibit the expression of *FT* and delay flowering.

Phytohormones involve in flowering time regulation by interacting with or integrating into the other flowering time pathways at various levels. The different phytohormone pathways regulate multiple aspects of plants growth and development, the complex crosstalk between the pathways at various levels and their interaction with the flowering pathways ensure the coordinated response to flowering inducing signals. In addition, phytohormones such as ABA, JA are activated in response to external stress, they may serve as a link between flowering time regulation and external stress, such as drought stress, salt stress, herbivore predation and pathogen infection.

## **1.2 Quantitative trait loci analysis (QTL) and its application on natural variation studies**

### **1.2.1 QTL analysis**

Arabidopsis natural populations have broad geographic distribution across the northern atmosphere, ranging from 68°N (North Scandinavia) to 0° (mountains of Tanzania and Kenya) (Koornneef et al., 2004). In order to adapt their local environment, Arabidopsis populations display extensive allelic variation in many loci controlling various traits, including flowering time.

Flowering time in Arabidopsis is determined by a complex network of genes integrating the internal physiological signals and the external environmental signals. As a result, Flowering time exhibits a continuous distribution across the populations and is a quantitative trait. All the loci, which harbor flowering time genes and contribute to flowering time variations, are referred to as quantitative trait loci (QTL) (Alonso-Blanco et al., 2009). The analysis of the genetic basis of quantitative traits in populations derived from experimental crosses is known as QTL analysis. QTL analysis evaluates the statistical significance of associations between phenotypic variation and specific alleles at and in between marker loci. The QTL approach is especially useful when analyzing traits under the control of multiple genes, such as flowering time. The number of QTL and their interaction reflect the genetic architecture, the contribution of each locus in the variance of the trait is also calculated in the QTL analysis.

The ultimate goal of flowering QTL analysis is to identify genes harboring polymorphisms that cause the flowering time variation. The understanding of such natural variation can provide insights in adaptation mechanisms and the evolutionary history of Arabidopsis (Agrena et al., 2013; Grillo et al., 2013).

### **1.2.2 Recombinant inbred line (RIL) as mapping population**

Two things are important to conduct a QTL analysis: a mapping population and the genotype of each line in the mapping population. RILs are often used in QTL analysis as a mapping population. RIL population derive from a cross between two accessions; the F<sub>2</sub> generation is propagated by selfing and through single seed descendant until F<sub>7</sub>-F<sub>10</sub> generation. The heterozygosity across genome is halved in each generation, eventually each line in a F<sub>9</sub> or F<sub>10</sub> RIL

population has a nearly homozygous genome, which is comprised of the parental genomes in a mosaic pattern. The homozygosity allows stable inheritance of the genome, so the exact genetic material can be analyzed multiple times for different traits under different conditions. In addition, the genotype at markers across the genome is stable, so the RIL population does not need to be genotyped in every QTL analysis.

Near-isogenic lines (NIL) is another kind of stable mapping populations. In contrast to RILs, NILs carry only a single locus from one parent in the genomic background of the other parent. NILs are derived from repeated backcrossing to the background parent and extensive genotyping to exclude introgression outside the desired genomic region. Single Introgression in NILs may allow isolation of a single quantitative trait locus, and thus are often employed to validate the effect of a QTL detected in a RIL population. Depending on the population size, marker density and recombination sites, a validated QTL often contains numerous genes - the process to pinpoint the causal gene(s) of the QTL is called fine mapping.

### **1.2.3 Fine mapping and confirmation of candidate genes underlying a QTL**

Fine mapping narrows down the size of QTL by searching for lines with recombinant events within the QTL region and selecting lines exhibiting co segregation of phenotype and genotype. Ideally, fine mapping can narrow down the QTL to a single gene locus; but in many cases a number of genes remain in the QTL in the segregating recombinant lines. A candidate gene can be selected based on the presence of polymorphisms and known descriptions (Alonso-Blanco et al., 2005). A final proof for identification of a gene as causal to the QTL effect is often a complementation or allelism tests (Salvi and Tuberosa, 2005; Yano, 2001). If a mutant line deficient in the candidate gene differs from the wild type plant in the QTL trait, alleles of the candidate genes can be cloned and transformed into the background. If the transgenic plants carrying the two alleles are different in the QTL traits, the allelic variation is confirmed. Another method is to cross line carrying one allele of the candidate gene to parent carrying a different allele. A new QTL analysis can be conducted in the F<sub>2</sub> population raised from such cross and if the QTL persist, the candidate gene is confirmed (Gazzani et al., 2003).

The research in this thesis is based on the flowering time QTL analysis conducted in the RIL population derived from a cross between *Ler* and An-1 (Tisne et al., 2010). I identified and



characterized the underlying genes of two flowering QTL detected in this analysis. I identified the causal gene of the QTL at the top of chromosome 4 as the flowering time gene *FRIGIDA* and confirmed allele variation between the *Ler* allele and An-1 allele by transforming and expressing them in Col-0 background. I extended the study to more natural alleles of *FRI* and transformed 21 additional alleles in Col-0 background. The results showed that the *Ler* allele of *FRI* is semi-functional, in contrast to common belief; and a novel non-functional *FRI* allele from Wa-1 accession was discovered. The Wa-1 allele carries an amino acid mutation which destabilizes the FRI protein and thus abolishes its function.

In the second part of the thesis, I confirmed the flowering time QTL at chromosome 3 (QTL3) in three NIL populations segregating at QTL3 region between *Ler* and An-1, Sha or Eri. Fine mapping of QTL3 was conducted in the *Ler* x Eri line and narrow down the QTL region to a single gene locus: *VIP HOMOLOG 2 (VIH2)*. I conducted further genetic and biochemical analysis to characterize this novel flowering time gene.

## Chapter 2

---

# Functional Analysis of the Landsberg *erecta* and Wa-1 alleles of *FRIGIDA*

### Summary

Most of the natural variation in flowering time in *Arabidopsis thaliana* can be attributed to allelic variation at the gene *FRIGIDA* (*FRI*, AT4G00650), which activates expression of the floral repressor *FLOWERING LOCUS C* (*FLC*, AT5G10140). Usually, late-flowering accessions carry functional *FRI* alleles that encode full-length proteins, whereas early flowering accessions contain non-functional alleles that encode truncated version of *FRI* proteins. The two most frequent alleles found in early flowering accessions are the ones present in the common lab strains Columbia (*FRI-Col*) and Landsberg *erecta* (*FRI-Ler*), that are considered null alleles due to a premature stop codon and a deletion of the start codon respectively.

During my thesis, I analyzed flowering time data from various sources and found evidences that *FRI-Ler* retains some functionality with respect to the truly null *FRI-Col* allele, challenging the current functional classification for *FRI*. To test this hypothesis, I functionally characterized a number of *FRI* alleles from different *Arabidopsis* accessions by cloning and transforming them in a common genetic background and characterizing the molecular and physiological variation induced in the resulting lines. I have found that, in contrast to common belief, *FRI-Ler* is able to upregulate the expression of *FLC* and delay flowering time; though its effect is weaker compared to the fully functional *FRI* alleles due to its reduced expression levels. I have also identified a novel non-functional allele present in the accession Wa-1 that encodes a full-length *FRI* protein whose stability is affected by a single amino acid mutation.

In summary, this chapter of my thesis increases the accuracy of the current functional classification for *FRI*, one of the largest contributors to the natural variation in flowering time found in *Arabidopsis*.

## 2.1 INTRODUCTION

Two main life history strategies have been observed in natural populations of *Arabidopsis*: winter annual plants germinate in autumn, survive winter as a rosette, flower and set seeds in the following spring; whereas summer-annual plants germinate in spring or summer and finish one or several reproduction cycles in the same year (Koornneef et al., 2004; Pigliucci et al., 2003). These distinct strategies are mainly dependent on the allelic state at the vernalization pathway genes *FRIGIDA* (*FRI*) and *FLOWERING LOCUS C* (*FLC*), which act epistatically to regulate flowering time (Johanson et al., 2000; Michaels and Amasino, 1999).

*FLC* is a major flowering repressor whose expression is epigenetically silenced at the appropriate time to allow flowering. The coding sequence of *FLC* is highly conserved among *Arabidopsis* accessions (Li et al., 2014). Null alleles of *FLC* due to protein truncation or expression suppression have been reported, but are rare (Werner et al., 2005). Most polymorphisms in *FLC* locus are present in the non-coding sequences and are associated with different vernalization requirement in terms of length and temperature (Li et al., 2014). Nevertheless, the effect of *FLC* variation can only be observed in the presence of a functional allele of *FRI* (Caicedo et al., 2004).

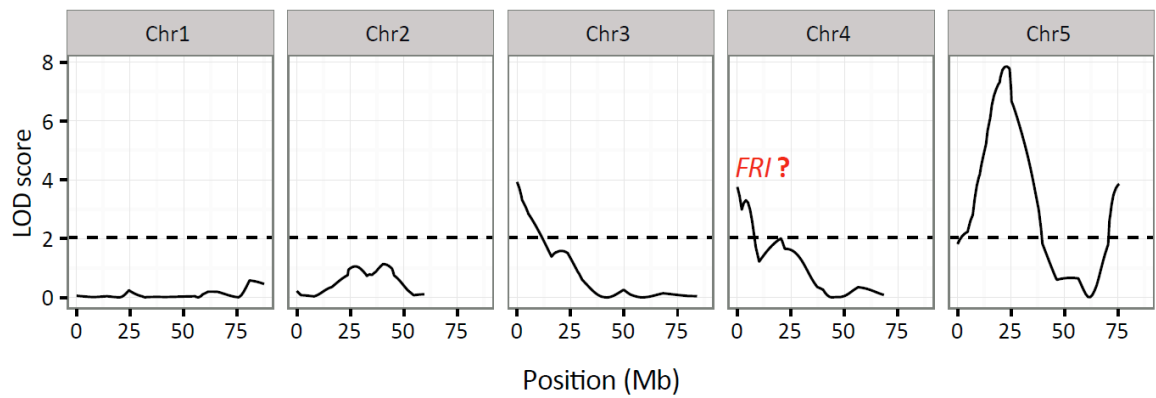
*FRI* is one of the most important determinants of flowering time in *Arabidopsis* accessions, especially when the plants are grown in the greenhouse condition without vernalization. Depending on the study, up to 70% of the variation in flowering time was attributed to allelic variation at *FRI* (Shindo et al., 2005).

Extensive nucleotide diversity has been observed in the *FRI* locus in nature, especially in the coding region. Comparison of the *FRI* protein sequence from *Arabidopsis* with that of closely related species *A. lyrata* reveals that the ancestral allele of *FRI* is as the sequence in the accession H51 (*FRI-H51*) (Le Corre et al., 2002). *FRI-H51* encodes a protein of 609 amino acids (aa) on three exons and it induces *FLC* expression, thus delaying flowering time in the absence of vernalization (Johanson et al., 2000). A considerable number of additional *FRI* alleles have been identified within accessions from a wide range of latitudes (Hagenblad et al., 2004; Lempe et al., 2005; Shindo et al., 2005) or more restricted geographic regions (Le Corre et al., 2002; Mendez-Vigo et al., 2011). Accessions carrying amino acid mutations in the *FRI-H51* allele, especially in the first exon, are widespread in nature. Most of these accessions are winter annual plants and require vernalization to

flower, regardless of their mutations. The function of several such *FRI* alleles has been confirmed by genetic transformation or crossing to an accession carrying non-functional *FRI* alleles (Gazzani et al., 2003). Consequently, it is widely accepted that all *FRI* alleles encoding full-length protein are functional, regardless of amino acid point mutations. For simplicity, all *FRI* alleles encode full-length proteins will be referred to as *FRI-fl* in the thesis. On the other hand, accessions that do not require vernalization for flowering usually carry non-functional alleles of *FRI*, which contain deleterious mutations such as premature stop codons and frame shifts resulting from point mutation or small insertions and deletions. The two most widespread *FRI* alleles in early flowering accessions are found in the common lab strain Columbia (Col-0) and Landsberg *erecta* (*Ler*). *FRI-Col* contains a 16-bp-deletion, which causes a premature stop codon after 314 aa, hampering the function of the protein. *FRI-Ler* contains a 376-bp-deletion combined with a 31-bp-insertion in the promoter region that disrupts the original translational start. It was hypothesized that an out-of-frame protein of 42 aa could be translated due to an alternative start codon in the insertion sequence (Johanson et al., 2000).

Because of its variability in nature and its major effect in flowering time, the region of chromosome 4 where *FRI* is located often appears as significant in QTL analyses on the flowering time of Arabidopsis (el-Lithy et al., 2006; Huang et al., 2011; Kover et al., 2009; Simon et al., 2008). One of such QTL studies was performed on a recombinant inbred line (RIL) population derived from a cross between the accessions An-1 and *Ler* (Tisne et al., 2010) (Fig. 2-1). *FRI* is an obvious candidate underlying the chromosome 4 locus; but this would not be possible in the presence of two non-functional alleles: the *FRI-Ler* allele and the *FRI-An-1*, which is described to carry a premature stop codon at 410 aa (Shindo et al., 2005). This observation led us to investigate whether any of these two alleles is not a true null allele by using bioinformatics, as well as molecular and physiological assays.

In this study, we characterized natural variation present in *FRI* in Arabidopsis in detail. We provide the first robust evidence that the *Ler* allele of *FRI* is functional, albeit its low expression, in contrast to the real null *FRI* allele in Col-0. Furthermore, we find that the Wa-1 accession that contains a full-length protein carries a point mutation that renders the protein non-functional. Characterization of this mutation provides us with insight about the structure of the *FRI* protein.



**Figure 2-1. QTL study on flowering time of a RIL population derived from a cross between An-1 and *Ler* accessions (Tisne et al., 2010).**

The 118 RILs were grown 12 hour light condition in the PHENOPSIS automated phenotyping platform (Granier et al., 2006) flowering time was recorded as the time between germination and emergence of the first flower buds.

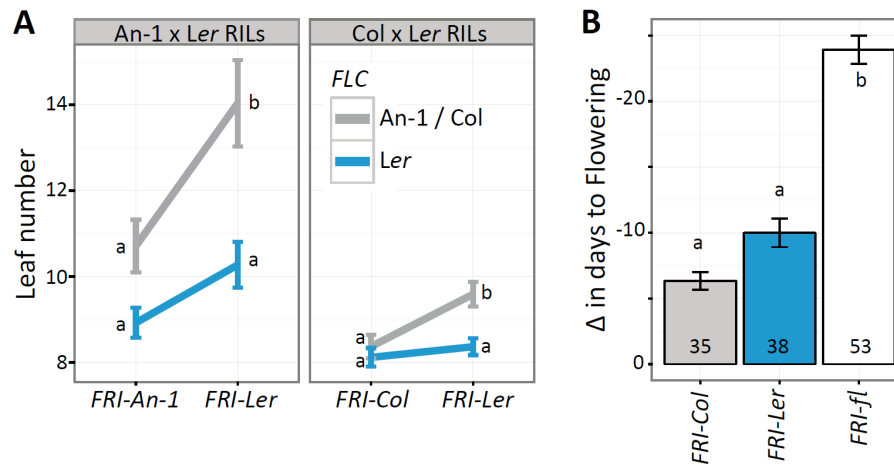
## 2.2 RESULTS

### 2.2.1 Analysis of published data suggest *FRI-Ler* functionality

In order to study the functionality of the *FRI-Ler* allele, we analyzed available flowering time data from two RIL populations derived from crosses between *Ler* and Col-0 or *Ler* and An-1, two accessions containing non-functional *FRI* alleles (Clarke et al., 1995; Tisne et al., 2008; Tisne et al., 2010). ). *FRI-Col* contains a 16-bp-deletion, resulting in a premature stop codon after 314 aa, hampering the function of the protein (Johanson et al., 2000); whereas, *FRI-An-1* carries a premature stop codon at 410 aa (Shindo et al., 2005). Because the effect of *FRI* on flowering time depends on the alleles at *FLC* (Michaels and Amasino, 1999), we considered the genotype of individual RILs at both loci. While An-1 and Col-0 carry a functional *FLC* allele, *Ler* contains a weak allele due to the insertion of a 1.2kb transposable element in its first intron (Michaels et al., 2003). As shown in Figure 2-2A, plants carrying the *FRI-Ler* allele flower significantly later than those carrying the An-1 or Col-0 alleles; although only in the presence of a functional *FLC* allele. This epistatic interaction, though not statistically significant (two-way ANOVA, interaction  $p = 0.114$  in An-1 x *Ler* RILs and  $p = 0.107$  in Col-0 x *Ler* RILs), suggests the existence of a functional *FRI* allele in *Ler*, as no other gene located in this region of chromosome 4 has been shown to delay flowering time through interactions with a locus in the region of chromosome 5 containing *FLC*.

The strong association between the function of *FRI* and the requirement for vernalization prompted us to investigate the effect of *FRI-Ler* on the vernalization response in a published dataset containing flowering time measurements and *FRI* genotypes for a worldwide set of 126 *Arabidopsis* accessions (Lempe et al., 2005). Accessions carrying the *FRI-Ler* allele show an intermediate vernalization response with respect to accessions carrying the *FRI-fl* alleles or the *FRI-Col* alleles (Fig. 2-2B); although the response is only significantly higher than that of the accessions with *FRI-Col* alleles when the accessions with *FRI-fl* were excluded from the analysis (t-test  $p = 0.029$ ).

Taken together, our results strongly suggest that the *FRI* allele present in *Ler* delays flowering and induces the vernalization requirement, functionally opposed to the true null Col-0 allele (Johanson et al. 2000). We focused the following studies in comparing the *FRI-Ler* allele to *FRI-Col*.



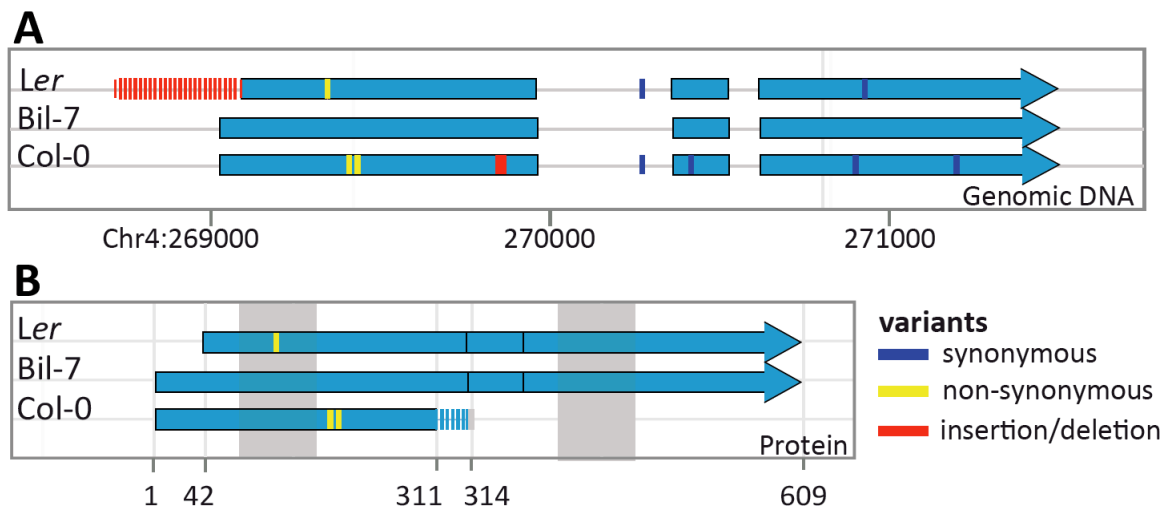
**Figure 2-2. Evidence of *FRI-Ler* functionality in published data.**

**(A)** Leaf number at flowering from individuals in the An-1 x Ler or Col x Ler RIL populations grouped by their genotype at *FRI* and *FLC*. **(B)** Vernalization response of Arabidopsis natural accessions measured as the reduction of days to flowering in plants vernalized for 5 weeks at 4°C compared to un-vernalized plants. Accessions are grouped by their *FRI* alleles. The number at the bottom of each bar indicates the number of accessions in each group. The error bars in (A) and (B) correspond to the standard error of the mean. Letters at each bar or point represent significance groups as determined by the Tukey HSD test.

### 2.2.2 *FRI-Ler* may encode a N-terminus truncated functional protein

We analyzed the sequence of the three common *FRI* alleles in nature, the *FRI-Col* allele, the *FRI-Ler* allele and the fully functional allele present in accession H51 (Johanson et al. 2000). To do this, we amplified by PCR and sequenced the *FRI* genomic region (including about 1.5 kbp promoter and 0.6 kbp terminator sequence) in Col-0, *Ler* and Bil-7, which contains the H51 allele of *FRI* (Shindo et al., 2005). Sanger sequencing results confirmed the presence of the expected allele in the three accessions (Fig. 2-3A). As mentioned in the introduction, the H51 allele is predicted to encode a fully functional 609 aa protein (GenBank: AAG23415) while the Col-0 allele yields a 314 aa protein (GenBank: AEE81913), truncated due to a premature stop codon resulting from a 16-bp-deletion in exon 1. In the case of *FRI-Ler*, a 376 bp deletion combined with a 31 bp insertion removes the translational start, but creates a new, out-of-frame start codon that is

predicted to yield a 41 aa protein (Johanson et al., 2000). Interestingly, in addition to the out-of-frame start codon, *FRI-Ler* contains an in-frame start codon downstream of the original start. Translation from this start codon would result in a protein missing the first 42 aa of the N-terminus and carrying one conservative amino acid change from leucine to isoleucine at position 79 (L79I) in the first coiled-coil domain compared to *FRI-Bil-7* (Fig. 2-3B). It has been shown that the N-terminus truncated FRI-H51 protein lacking the first 118 aa is partially functional (Risk et al., 2010). We hypothesized that if the transcription and translation occurred from the downstream in-frame start codon of the *FRI-Ler* allele, the resulting protein could be partially functional or even as fully-functional. To test this hypothesis, we analyzed RNA-seq reads from seven Arabidopsis accessions containing the *FRI-Ler* allele, and found that the reads covered the complete C-terminus domain (Gan et al., 2011). This raised the question of whether this long *FRI-Ler* transcript is translated into a functional protein.



**Figure 2-3. Genomic and protein sequence comparison between *FRI-Ler* and *FRI-Col* to *FRI-Bil-7*.**

**(A)** Exonic organization for *FRI-Col*, *FRI-Ler* and *FRI-Bil-7*. Red regions indicate the indels characteristic for *FRI-Col* and *FRI-Ler*. **(B)** Predicted proteins for all three alleles shown in (A). The dashed region in *FRI-Col* represents a shift in the ORF caused by the 16 bp deletion. Shaded regions in the background indicate predicted coiled-coil domains.



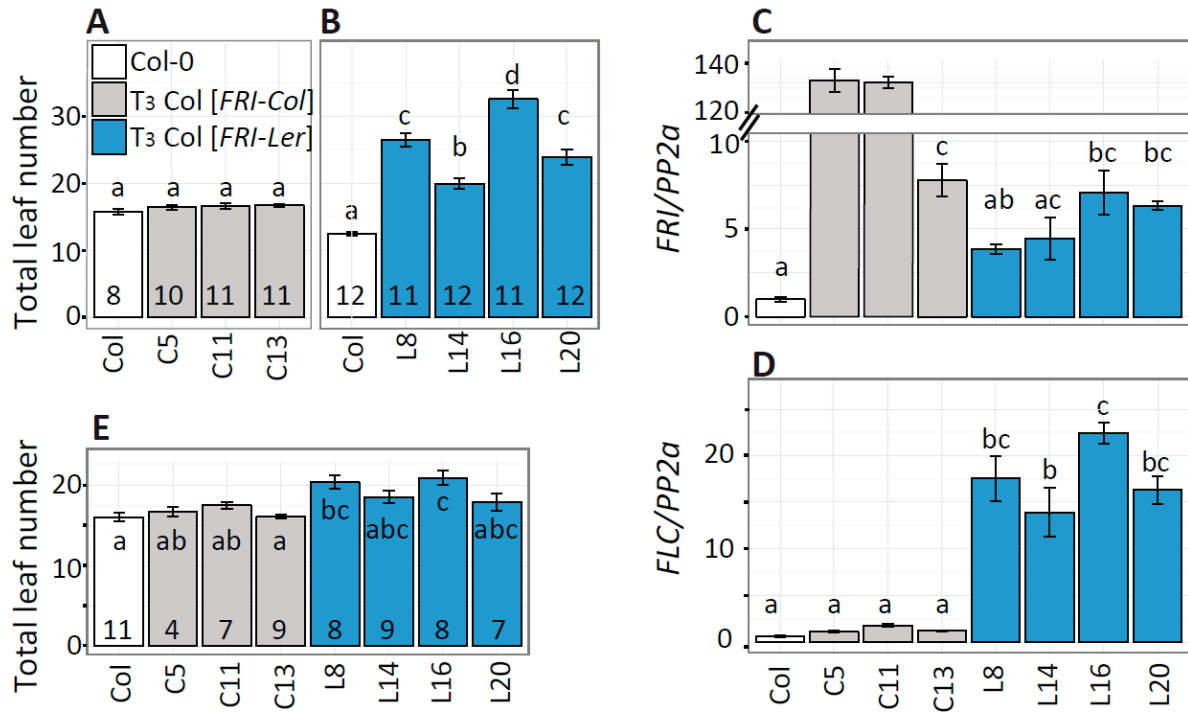
### **2.2.3 Functionality of *FRI-Ler* is confirmed in transgenic plants**

To confirm the functionality of *FRI-Ler* allele, we compared the *FRI-Ler* and *FRI-Col* alleles in transgenic plants. The genomic region of *FRI* was cloned from the *Ler* and Col-0 accessions, making use of the BamHI and EcoRI restriction sites at -1372 bp and +2691 bp respectively relative to the annotated *FRI* start codon in TAIR10. Col-0 was chosen as the transgenic background because our previous results suggested that *FRI-Col* is indeed a loss-of-function allele. In addition, Col-0 carries a strong *FLC* allele that is activated in the presence of a functional *FRI* allele (Johanson et al., 2000).

First, we confirmed the lack of function of the *FRI-Col* allele. T<sub>3</sub> lines homozygous for the *FRI-Col* transgene (Col-0 [*FRI-Col*]) were grown under long day conditions in the greenhouse. None of the transgenic lines showed flowering time significantly different from wild type (Fig. 2-4A). In a consecutive experiment in similar conditions we observed that homozygous T<sub>3</sub> transgenic lines containing the *FRI-Ler* transgene (Col-0 [*FRI-Ler*]) flowered significant later than the Col-0 wild type (Fig. 2-4B), confirming the functionality of the *FRI-Ler* protein.

The expression of the *FRI* transgene was confirmed by quantitative real time PCR (qRT-PCR). All the transgenic lines showed higher *FRI* expression compared to the endogenous expression of *FRI-Col* in the wild type, indicating the presence of additional transcripts from the transgene (Fig.2-4C). Expression of the transgene varied among different lines, likely due to variation in the position of insertion and copy number of the transgene. It is worth to note that two of the three Col-0 [*FRI-Col*] lines showed extremely high levels of transgene expression, which may lead to high level of *FRI-Col* protein accumulation or silencing of both the transgene and the endogenous *FRI* (Napoli et al., 1990). Nevertheless, independently of the level of expression of the transgene, none of the Col-0 [*FRI-Col*] lines flowered later than the wild type plants; further proving the lack of function of the *FRI-Col* allele. We also measured the expression of *FLC* in the transgenic plants by qRT-PCR because *FRI* delays flowering time through up-regulation of *FLC* (Johanson et al., 2000). As expected, elevated *FLC* expression was observed in all *FRI-Ler* transgenic lines but not in the *FRI-Col* lines (Fig. 2-4D). The flowering time of the *FRI-Ler* transgenic lines reflected precisely their *FLC* expression level. In addition, a four-week-vernalization treatment significantly reduced leaf number in Col-0 [*FRI-Ler*] lines, but not in Col-0 [*FRI-Col*] lines or Col-0 wild type (two-way ANOVA,  $p < 0.001$  for the interaction between

vernalization and genotypes, comparing Fig. 2-4E to A and B). Altogether, our results demonstrate that the *FRI-Ler* allele, but not the *FRI-Col* allele, is able to delay flowering through up-regulation of *FLC*.



**Figure 2-4. Characterization of transgenic lines carrying the *FRI-Col* or *FRI-Ler* allele.**

**(A, B)** Flowering time expressed as total leaf number from homozygous T<sub>3</sub> single insertion lines containing a *FRI-Col* (A) or *FRI-Ler* (B) transgene in comparison to the transgenic background Col-0. All plants were grown under long day condition in a greenhouse. **(C, D)** Expression of *FRI* (C) and *FLC* (D) for the same genotypes as in (A, B). In this experiment, aerial parts were collected from 10-day-old seedlings grown in long day conditions in an environmental chamber. Three biological replicates were used for each genotype. Expression was normalized to the expression of *PP2a*. **(E)** Flowering time after vernalization quantified as total leaf number for the same lines as in (A-D). Plants were vernalized for four weeks at 4°C and subsequently grown under long day conditions. Error bars represent the standard error of the mean. The number in each bar indicates the number of individual plants per line analyzed. Letters in each bar represent significance groups as determined by the Tukey HSD test.

#### 2.2.4 The *FRI-Ler* allele presents expression defects due to cis-regulation

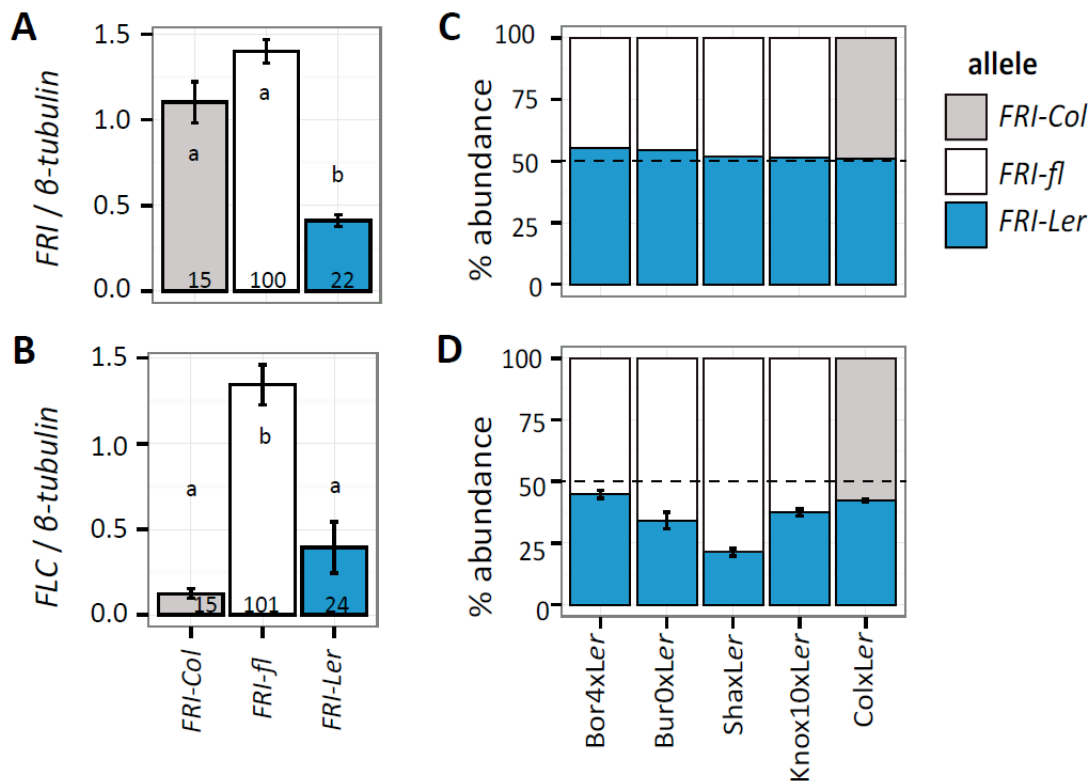
The functionality of the *FRI-Ler* allele is contrasting with its frequent presence among early flowering accessions (Lempe et al., 2005; Shindo et al., 2005). Our analysis also showed that although accessions containing the *FRI-Ler* allele have stronger vernalization responses than accessions containing *FRI-Col* allele, they do not reach the levels observed in accessions with *FRI-fl* alleles that encode full-length proteins (Fig. 2-2B). Previous works have shown low expression of *FRI* in *Ler* accession (Johanson et al., 2000; Shindo et al., 2005), leading us to investigate whether *FRI-Ler* is transcriptionally inhibited and thus causes the early flowering of the natural accessions carrying this allele.

We analyzed the expression level of *FRI* and its downstream target *FLC* in 137 Arabidopsis accessions grouped by their *FRI* alleles (Atwell et al., 2010). Expression of *FRI* is significantly lower in accessions carrying *FRI-Ler* than those carrying the *FRI-Col* or *FRI-fl* (Fig. 2-5A). However, the reduced expression of *FRI-Ler* was still able to moderately up-regulate the expression of *FLC* in comparison to lines expressing the *FRI-Col* allele, though the difference in *LFC* expression is not significant between the *FRI-Ler* and *FRI-Col* groups ( $p = 0.079$  in one-way ANOVA); possibly due to the large variation among the accessions containing *FRI-Ler* alleles (Fig. 2-5B). In the same study (Atwell et al., 2010), a GWAS analysis on *FRI* expression using 164 natural accessions revealed a single locus at *FRI*, indicating that the transcriptional inhibition at this locus is mediated in *cis*.

In order to confirm the *cis*-regulation of *FRI-Ler* expression, my colleague Malgorzata Rynhajllo measured allele specific *FRI* expression using pyrosequencing in F<sub>1</sub> hybrids from crosses between *Ler* and other accessions carrying either *FRI-Col* (accession Col-0) or *FRI-fl* (accessions Bor-4, Bur-0, Sha or Knox-10) alleles. Because expression of both parental alleles in a hybrid cell is controlled by the same set of regulators, the relative abundance of each allele in the hybrid should be equaled out if *trans*-regulators cause the differences between accessions. In contrast, if the expression differences between accessions are caused by mutations in *cis*, the differences in the expression of the alleles in the hybrid will be maintained (Wittkopp et al., 2004).

We first ensured that the primers used in the assay had no preference for any of the two alleles by measuring the relative abundance of each parental allele in genomic DNA from all F<sub>1</sub> individuals used. In all cases, the relative abundance of the alleles in the hybrids was close to 50%,

confirming the feasibility of the assay (Fig. 2-5C). Analysis of cDNA from the same individuals revealed lower relative abundance of *FRI-Ler* than any other alleles, independently of whether they were *FRI-fl* alleles (in hybrids between *Ler* and Bor-4, Bur-0, Sha or Knox-10) or *FRI-Col* alleles (in hybrids between *Ler* and Col-0) (Fig. 2-5D). These differences confirmed that, at least in part, the early flowering of accessions carrying the *FRI-Ler* allele is due to a *cis*-regulated reduction in *FRI* expression, most likely due to the deletion in the promoter region. Interestingly, differences between expression of *FRI-Ler* and *FRI-fl* varied widely among the hybrids (Fig. 2-5D), and were in general smaller than the differences found in the accessions (Fig. 2-5A). The variation in transcription activity of the different *FRI-fl* alleles in the hybrids could be explained by polymorphisms in their promoter sequences, but further studies will be required to determine the precise cause.



**Figure 2-5. Expression of *FRI-Ler* is lower in Arabidopsis accessions and in F<sub>1</sub> hybrids.**

**(A, B)** 137 Arabidopsis accessions were grown under long day conditions in a greenhouse; sampled after four weeks and expression levels of *FRI* (A) and *FLC* (B) were quantified from northern hybridizations (Atwell et al., 2010). Letters in each bar indicate the significance groups as determined by the Tukey HSD test. The number in each bar indicates the number of accessions in

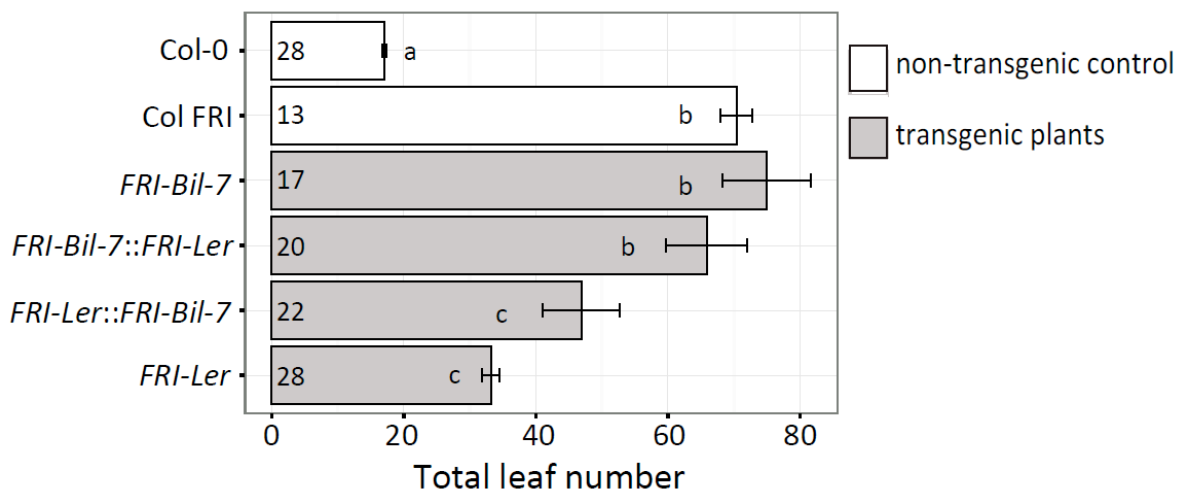
the group. **(C, D)** Abundance of specific SNP between *Ler* and other accessions were measured using pyrosequencing in genomic DNA (C) or in cDNA (D) of F<sub>1</sub> hybrids. Leaf material for hybrids involving Bor-4, Bur-0, Knox-10 or Sha was collected at the time of bolting from plants growing under 12-hour photoperiods in an environmental chamber. Leaf material from hybrids involving Col-0 was collected from 10-day-old seedlings grown in the greenhouse under long day conditions. Error bars indicate the standard error of the mean.

### **2.2.5 Lack of effect of coding polymorphisms in *FRI-Ler***

As mentioned previously, the predicted *FRI-Ler* protein lacks 42 aa at N-terminus and carries a conservative L79I mutation in comparison to *FRI-Bil-7* (Fig. 2-3B). To test whether these polymorphisms have any effect on the function of the *FRI* protein, we generated transgenic lines carrying promoter swap constructs, where the coding region of the *FRI-Bil-7* and *FRI-Ler* alleles was placed downstream of the native promoter of the other allele (*FRI-Bil-7::FRI-Ler* and *FRI-Ler::FRI-Bil-7*). As a positive control, we used plants carrying a construct with the *FRI-Bil-7* coding sequence under the control of its native promoter (*FRI-Bil-7*). T<sub>1</sub> transgenic plants carrying the above-mentioned constructs were grown under long day conditions. The previously characterized T<sub>3</sub> Col-0 [*FRI-Ler*] plants were grown as control, together with the Col-0 non-transgenic plants and a near isogenic line with a functional *FRI* allele from the accession Sf-2 introgressed into Col-0 (named Col *FRI*) (Michaels and Amasino, 1999).

As shown in Figure 2-6, transgenic plants carrying the *FRI-Bil-7* transgene flowered with the same number of leaves as the Col *FRI* control. Transgenic plants carrying constructs with the *FRI-Ler* promoter flowered significantly earlier than those carrying constructs with the *FRI-Bil-7* promoter; further confirming that the indel mutation in the *FRI-Ler* promoter region is the cause of the early flowering in accessions carrying the *FRI-Ler* allele. Interestingly, lines carrying the *FRI-Ler* coding region flowered earlier than lines carrying the *FRI-Bil-7* coding region when expressed under the *FRI-Ler* promoter ( $p = 0.053$  in t-test). This result suggests that the mutations in the *FRI-Ler* coding sequence promote post-transcriptionally inhibition or weaken the protein with respect to the *FRI-Bil-7* allele. From our analysis it is difficult to infer whether the N-terminal deletion or the leucine to isoleucine mutation is the cause of these differences; but the N-terminal deletion is most likely the cause. The significant differences in flowering time between lines carrying the *FRI-Ler*

and *FRI-Bil-7* coding sequence disappeared ( $p = 0.594$  in t-test) when both alleles were expressed under the *FRI-Bil-7* promoter. A possible explanation for this could be that the higher expression level induced by the *FRI-Bil-7* promoter compensates the reduced function of the *FRI-Ler* allele. Taken together my results suggest that the *FRI-Ler* allele shows reduced function mainly due to strong transcriptional inhibition caused by the deletion at the transcription start region, but the 42 aa N-terminal truncation or the L79I mutations in the protein may further weaken the function of *FRI-Ler* protein when expressed at low level.



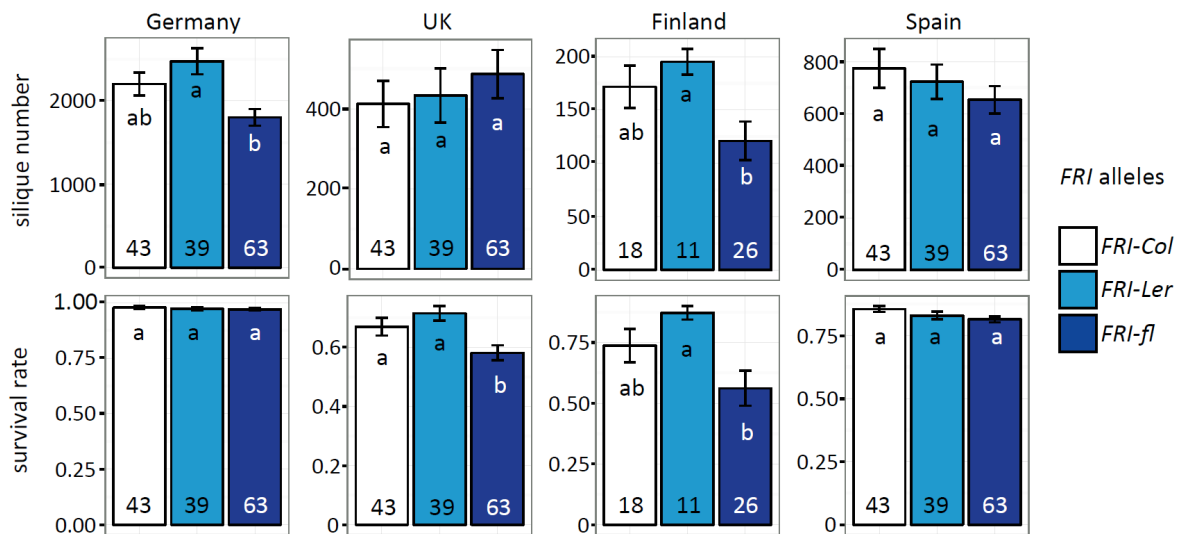
**Figure 2-6. Flowering time of transgenic lines carrying promoter-swap constructs involving the *FRI-Ler* and *FRI-Bil-7* alleles.**

Plants were grown under 16-hour light conditions in an environmental chamber. Total number of leaves was counted on the day when the first flower opened. Col-0 is the untransformed control and Col FRI is a near isogenic line carrying the functional *FRI* allele from the accession Sf-2. Numbers in each bar indicate the number of individual plants analyzed per genotype. Error bars represent the standard error of the mean. Letters in each bar represent significance groups as determined by the Tukey HSD test.

### 2.2.6 Fitness effect of the different *FRI* alleles

In the previous sections, we demonstrated that *FRI-Ler* is a semi-functional allele. This raises an obvious question: does the functionality of *FRI-Ler* affects the fitness of the accessions carrying it in comparison to *FRI-Col*? Although *FRI-Ler* accessions revealed a significant higher response to

vernalization than accessions with the *FRI-Col* allele (Fig. 2-2B), both groups are early flowering under greenhouse conditions (compared to accessions carrying the fully functional *FRI-fl* allele) and are unlikely to differ in their life history strategy, which is strongly associated with local adaptation and thus fitness. Nevertheless, we reanalyzed fitness-related phenotypes measured in a panel of accessions grown at four different field locations (Fournier-Level et al., 2011), grouping the accessions by our new functional classification of *FRI*. We observed small fitness differences in accessions carrying the *FRI-Col* and *FRI-Ler* alleles at various locations; however, these were not significant (Fig. 2-7). On the other hand, accessions containing *FRI-fl* alleles were often significantly different in fitness-related traits when compared with *FRI-Ler* and *FRI-Col*. Further experiments will be required to determine if differences between the *FRI-Ler* and *FRI-Col* allelic classes are relevant in specific environmental conditions (e.g. under stress).



**Figure 2-7. Fitness related traits of accessions carrying different *FRI* alleles growing in field.**

Accessions were sown in fall coinciding with the local germination. Survival plants were those successfully set seeds the following year and the silique number was the average value including individuals that did not survive (Fournier-Level et al., 2011). Accessions were grouped based on the published *FRI* genotypes: those with the characteristic deletions in *Col* or *Ler*, or encoding a full-length protein (Atwell et al., 2010; Lempe et al., 2005; Shindo et al., 2005; Wilczek et al., 2009). Numbers in each bar indicate the number of accessions analyzed per group. Error bars represent the standard error of the mean. Letters in each bar represent significance groups as determined by Tukey HSD test.

### 2.2.7 A new non-functional FRI allele: *FRI-Wa-1*

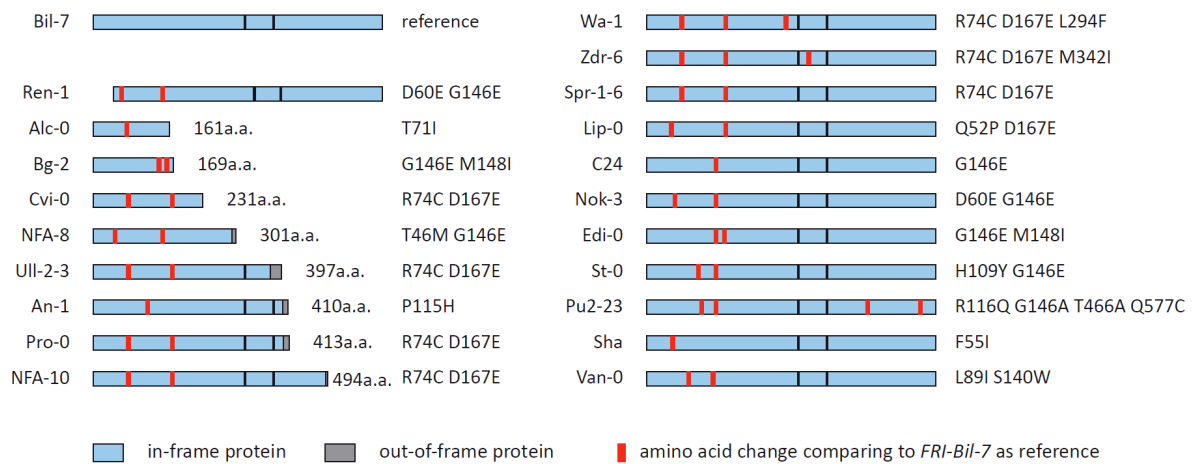
Our previous work showed that comparison of natural alleles of *FRI* through heterologous expression in a common genetic background is a suitable method to study functional differences. In this chapter I describe the analysis of a set of *FRI* alleles with different mutations. Specifically, we searched full-length *FRI* alleles that may not delay flowering in the Col-0 background, or truncated *FRI* alleles that may cause later flowering.

We selected twenty *FRI* alleles based on their sequence published by (Shindo et al., 2005). Eleven alleles were predicted to encode full-length protein and present full functionality. They carry amino acid mutations at various positions compared to FRI-Bil-7, most of the mutations occur in the N-terminus. Several mutations are very common among *FRI* alleles, such as G146E, M148I, R74D and D167E. Three alleles with mutations in the middle or in the C-terminus were selected in the study: Wa-1, Zdr-6 and Pu2-23. The other nine alleles are predicted to encode partial proteins of various lengths due to deleterious mutations (Fig. 2-8A). We also compared the genomic sequence of the promoter region in all the alleles. Single nucleotide polymorphisms and small deletions of 1 to 3 nucleotides are frequent among promoters. Major deletions were only found in accessions Ren-1 and St-0. *FRI-Ren-1* has a deletion from -337 to 47 bp relative to the position of start codon, but the protein is predicted to translate from the next in-frame start codon, resulting a N-terminus truncated like FRI-Ler. The deletion in St-0 is more distal, going from -540 to -470 bp, where no known regulator motif has been described.

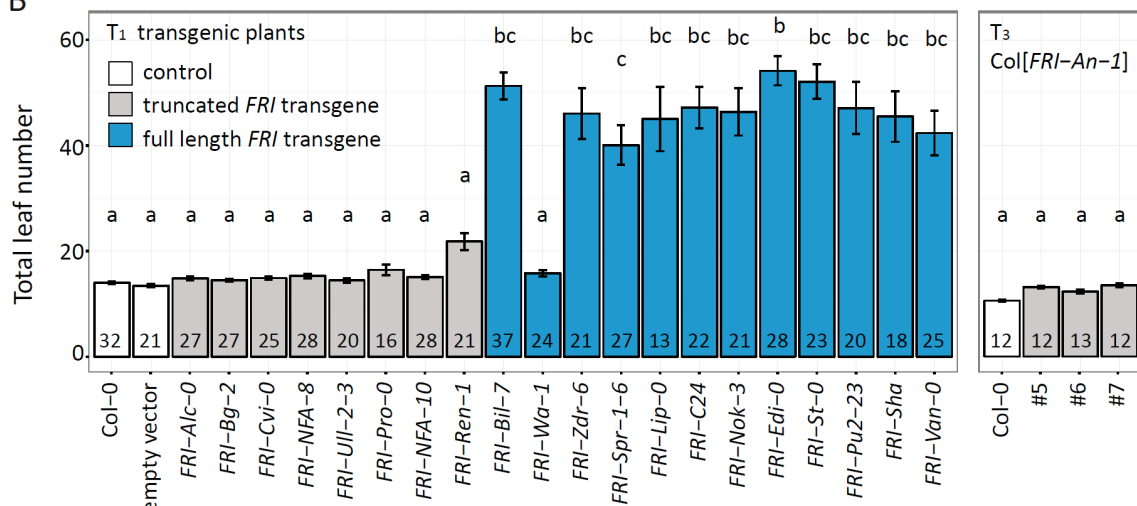
All the alleles were cloned with their respective native promoter regions and transformed into Col-0. We used a plasmid that contains a red fluorescence marker for selection of T<sub>1</sub> transgenic seeds under the microscope and without antibiotics. All transgenic seeds were germinated and grown under long day condition in a greenhouse. T<sub>3</sub> transgenic lines carrying *FRI-An-1* allele were used because seedlings require kanamycin selection on agar medium. All the lines carrying *FRI* alleles encoding C-terminus truncated proteins flowered at the same time as the Col-0 control plants (Fig. 2-8B). This result shows that the *FRI* alleles with C-terminal truncations of at least 115 amino acids are not functional.



A



B



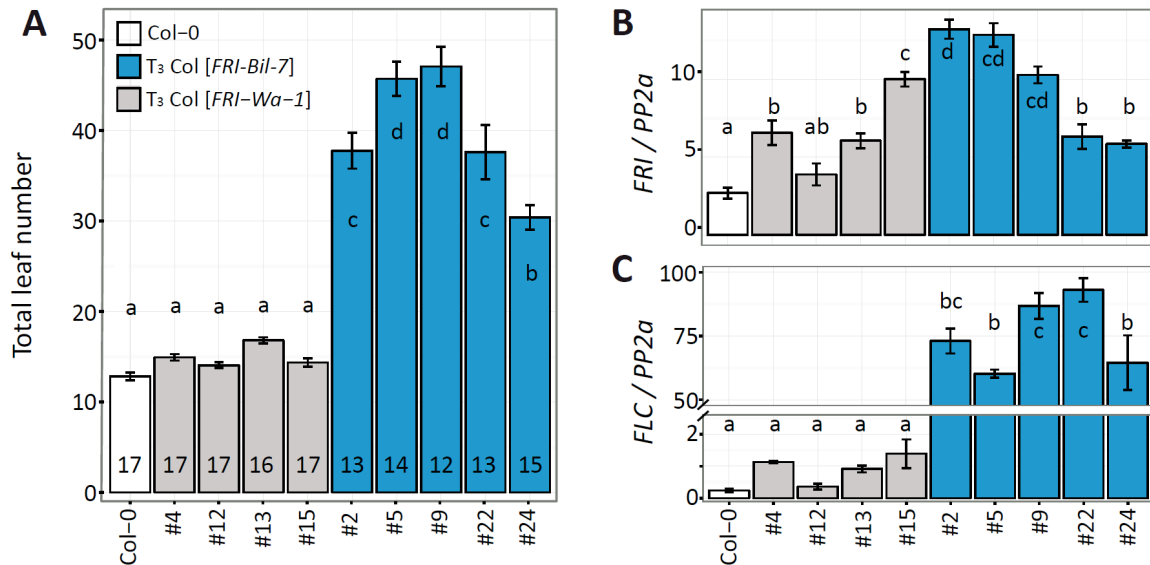
**Figure 2-8. Flowering time of transgenic plants carrying selected alleles of *FRI*.**

**(A)** Schematic representation of selected *FRI* proteins. Alleles on the left encode truncated proteins, those on the right encode full-length proteins. Red bars represent the positions of amino acids mutations relative to *FRI-Bil-7*. **(B)** Flowering time expressed as total leaf number in transgenic plants carrying the *FRI* alleles shown in (A). All plants were grown under long day conditions in the greenhouse. Error bars represent the standard error of the mean. The number in each bar indicates the number of individual plants per line analyzed. Letters on top of each bar represent significance groups as determined by the Tukey HSD test.

*FRI-Ren-1* was the only truncated *FRI* allele able to delay flowering in the transgenic plants. At flowering, the total leaf number was  $21.8 \pm 1.6$  in Col-0 [*FRI-Ren-1*], compared to  $14.0 \pm 0.2$  in Col-0 ( $p < 0.001$  in t-test between the two groups). Due to a deletion similar to *FRI-Ler* at the promoter region including the original translational start, the *FRI-Ren-1* protein is predicted to lack 42 aa from the N-terminus, as the *FRI-Ler* allele. The transgenic plants carrying *FRI-Ren-1* or *FRI-Ler* alleles showed similar flowering time phenotypes; despite that the *FRI-Ren-1* and *FRI-Ler* proteins have three amino acid mutations. This observation further confirmed the presence of a functional protein in accessions carrying deletions at the start codon.

Most *FRI* alleles encoding full-length proteins significantly delayed flowering when added to a Col-0 background (Fig. 2-8B). The transgenic line carrying the fully functional *FRI-Bil-7* allele flowered much later than Col-0 ( $51.3 \pm 2.6$  versus  $14.0 \pm 0.2$  total leaves respectively). The flowering time of most Col-0 [*FRI-fl*] lines were not significantly different from that of Col-0 [*FRI-Bil-7*], despite the differences in the promoter sequence or amino acid mutations among the alleles. However, the transgenic line carrying the *FRI-Wa-1* allele flowered earlier than Col-0 [*FRI-Bil-7*], but not significantly different from the control plants ( $15.8 \pm 0.6$  total leaves in Col-0 [*FRI-Wa-1*] versus  $14.0 \pm 0.2$  total leaves in Col-0), suggesting that the *Wa-1* allele of *FRI* carries a deleterious mutation.

We confirmed the loss of function in the *FRI-Wa-1* allele in  $T_3$  lines homozygous for a single copy of the *FRI-Wa-1* transgene (Col-0 [*FRI-Wa-1*]). Various independent  $T_3$  transgenic lines carrying *FRI-Bil-7* or *FRI-Wa-1* alleles were grown in the greenhouse under long day conditions. The effect of the two alleles on flowering time was similar to what was observed in the  $T_1$  generation (Fig. 2-9A). We confirmed the expression of both *FRI-Bil-7* and *FRI-Wa-1* transgenes in all the lines (Fig. 2-9B). Despite a large variation in expression levels among the transgenics, the generally high expression levels among all transgenic lines suggests that the reduced functionality of the *FRI-Wa-1* allele is not in principle associated to expression defects. Moreover, transgenics expressing the *FRI-Wa-1* allele did not present high expression levels of *FLC*, in contrast to the lines carrying the *FRI-Bil-7* allele (Fig. 2-9C). Because this includes lines with high and low *FRI-Wa-1* expression levels, this result confirms that the *FRI-Wa-1* allele must contain a coding polymorphism that renders the protein non-functional.



**Figure 2-9. Characterization of transgenic lines carrying the *FRI-Wa-1* or *FRI-Bil-7* alleles.**

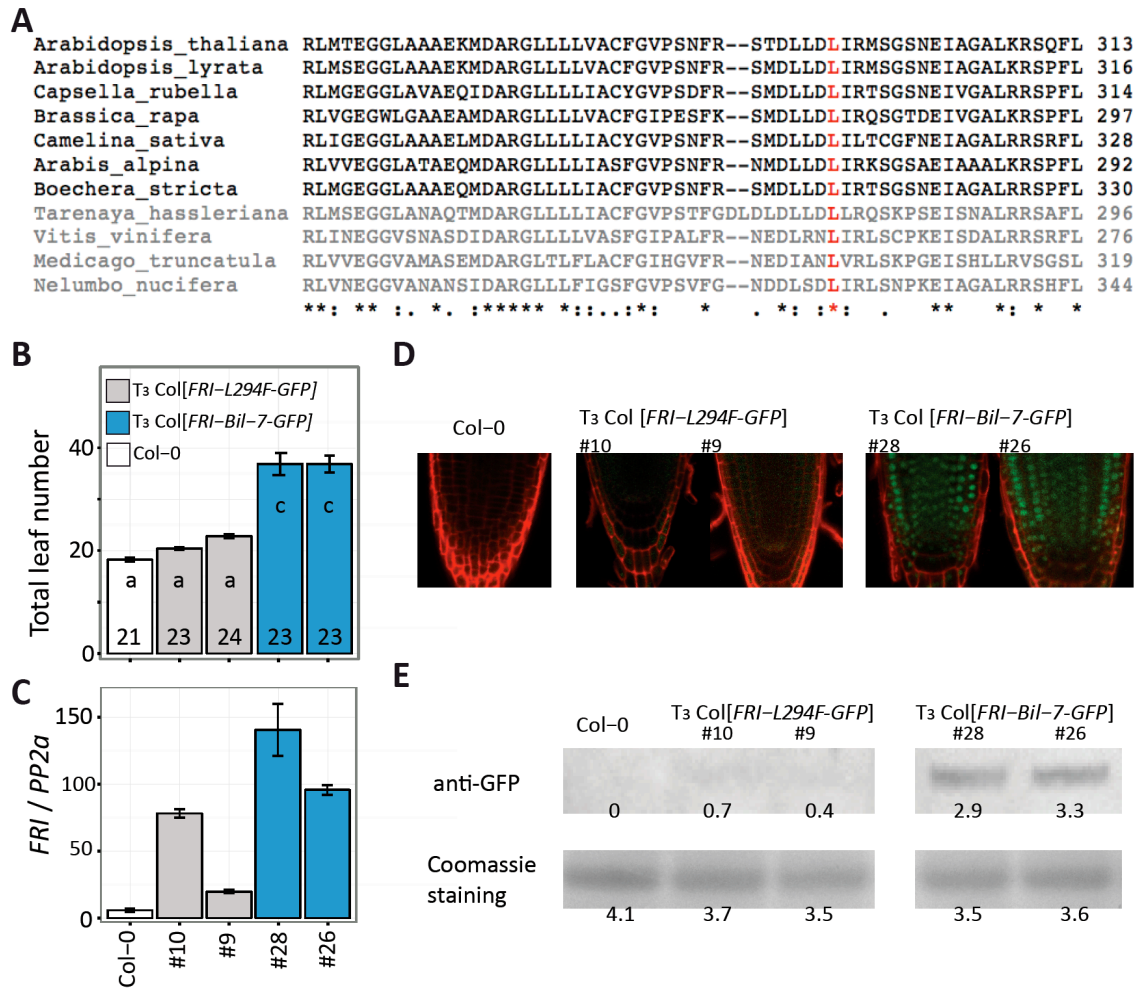
**(A)** Flowering time expressed as total leaf number from homozygous  $T_3$  single insertion lines containing a *FRI-Wa-1* or *FRI-Bil-7* transgene in comparison to the transgenic background Col-0. All plants were grown under long day conditions in the greenhouse. **(B, C)** Expression of *FRI* (B) and *FLC* (C) in the transgenic lines shown in (A). Aerial part was pooled from 10-day-old seedlings grown in long day conditions. Three biological replicates were used for each genotype. Expression was normalized to the expression of *PP2a*. Error bars represent the standard error of the mean. The number in each bar indicates the number of individual plants per line analyzed. Letters in each bar represent significance groups as determined by the Tukey HSD test.

### 2.2.8 Single amino acid mutation abolishes the function of *FRI-Wa-1* allele

We studied the protein sequence of the *FRI-Wa-1* allele as the putative cause of its loss of function. There are three amino acids differences between the *FRI-Wa-1* and *FRI-Bil-7* alleles (Fig. 2-8A). The first two mutation, R74C and D167E are common among functional *FRI* alleles, such as *FRI-Spr-1-6* and *FRI-Zdr-6*, and are unlikely the cause of the functional defect in *FRI-Wa-1*. However, the third mutation, L294F, is located in the conserved central domain of FRI and could have deleterious effects (Risk et al., 2010). We investigated the conservation of this position in the FRI protein across several *Brassicaceae* species and other less related species (Fig. 2-10A). The leucine mutated in FRI-Wa-1 is conserved in all proteins studied, suggesting that L294 could be an important amino acid for the function of the FRI protein. The change from leucine to phenylalanine

in FRI-Wa-1 is a non-conservative change because the isopropyl side chain is replaced by a much bulkier phenyl group. We next asked whether the L294F change alone is the cause of the functional defect in FRI-Wa-1 by constructing an artificial FRI allele (*FRI-L294F-GFP*) in which we added the L294F mutation and a C-terminal GFP tag to the FRI-Bil-7 protein. This construct and its control (*FRI-Bil-7-GFP*) were transformed into Col-0 plants. Homozygous T<sub>3</sub> transgenic plants carrying single insertion of the transgene were selected and grown in the greenhouse under long day conditions. Lines carrying the *FRI-L294F-GFP* allele flowered as early as the Col-0 control plants, whereas those carrying *FRI-Bil-7-GFP* flowered significantly later (Fig. 2-10B). As the expression of the transgene was high in all the transgenic lines (Fig. 2-10C), we concluded that the L294F mutation alone abolishes the function of FRI-Wa-1 protein.

In order to understand the effect of the L294 mutation on the FRI protein, we studied its cellular localization. For this we observed the GFP signals with confocal microscopy in root tips from transgenic seedlings grown on agarose medium. As reported in previous studies, we observed strong GFP signals in nuclei from Col [*FRI-Bil-7-GFP*] plants (Fig. 2-10D) (Kim et al., 2006). In contrast, only very weak GFP signals were observed in the Col [*FRI-L294F-GFP*] plants, evenly distributed in all cells without any recognizable pattern. Western blot results using the same plant material agreed with the differences in protein abundance observed under the microscope (Fig. 2-10E). The lack of accumulation of FRI protein with the L294F mutation allows us to conclude that this single mutation causes instability and degradation of the FRI protein and thus abolishes its function.



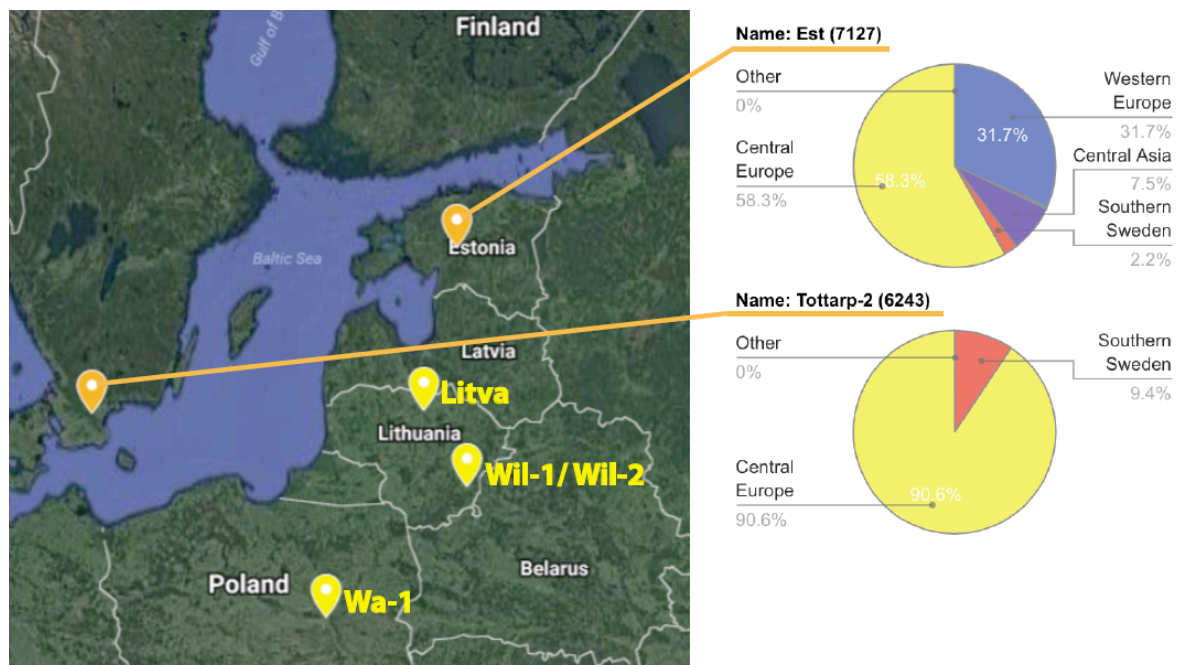
**Figure 2-10. Characterization of L294F amino acid mutation in FRI protein.**

**(A)** Alignment of FRI protein sequences from species in (in black) or outside (in grey) the *Brassicaceae* family. Alignments were produced with Clustal Omega. The leucine at position 294 mutated in *FRI-Wa-1* (in red) is conserved in all species. **(B)** Flowering time expressed as total leaf number from homozygous T<sub>3</sub> transgenic lines with *FRI-L294F-GFP* or *FRI-Bil-7-GFP* transgene in comparison to Col-0. All plants were grown under long day conditions in the greenhouse. The number in each bar indicates the number of individual plants per line analyzed. Letters in each bar represent significance groups as determined by Tukey HSD test. **(C)** Expression of *FRI* in the same genotypes shown in (B). Aerial part was pooled from 10-day-old seedlings grown in long day conditions. Three biological replicates were used per genotype. Expression was normalized to the expression of *PP2a*. **(D)** Confocal images of root tips of one-week-old seedlings grown on MS agar medium under long day condition. The root was treated with propidium iodide, which outlined the cells in red. GFP signals from the FRI proteins could only be observed in the nuclei of transgenic lines carrying the wild type *FRI-Bil-7* allele. One representative image from each genotype is shown.

(E) Western blot detection of FRI protein using anti-GFP antibodies. Plants from the same experiment shown in (D) were pooled for protein extraction.

### 2.2.9 Analysis of other accessions with *FRI-Wa-1* allele

Given the importance of the leucine at position 294, we set out to search other accessions with this mutation. There are only six accessions (Fig. 2-11) carrying such mutation among the set of 1135 accessions for which re-sequencing data is available (The 1001 Genomes Consortium, 2016). Four accessions (Wa-1, Wil-1, Wil-2 and Litva) belong to the group encompassing central Europe accessions while Tottarp-2 and Est are genetically admixed but share a large fraction of their genomes with central European accessions (Fig. 2-11). Our results implicate that the L294F mutation in *FRI* occurred once, most likely in a central Europe population, and then distributed to the modern accessions by migration or outcrossing.



**Figure 2-11. Collection sites of accessions carrying same *FRI* allele as *FRI-Wa-1*.**

Four accessions (in yellow) belong to the central Europe genetic group. Two accessions (in orange) share central Europe ancestry, but have been classified as admixed with other genetic groups. For these two admixed accessions, their genetic components are shown in the pie charts.

## **2.3 DISCUSSION**

The analysis of natural variation in flowering time in *Arabidopsis* focused on the *FRI* gene from very early stages, because of its major contribution to this trait (Burn et al., 1993; Clarke and Dean, 1994; Lee et al., 1993). Since its identification (Johanson et al., 2000), multiple studies surveyed the sequence and function variation of the *FRI* gene (Gazzani et al., 2003; Le Corre et al., 2002; Shindo et al., 2005; Werner et al., 2005). Until now, natural alleles of *FRI* were clustered into two distinct groups: alleles encoding full-length FRI proteins that have been classified as functional, and alleles with the *FRI-Ler* / *FRI-Col* deletions or other premature stop codons that were considered non-functional. Here we update this classification by reporting that *FRI-Ler*, the most common allele in the non-functional group, is partially functional, and that *FRI-Wa-1*, though encoding a full-length protein, is a loss-of-function allele.

Our work shows that the *FRI-Ler* allele, as well as other alleles with similar N-terminal truncations such as the one found in the Ren-1 accession, can be classified as a third intermediate functional group of *FRI* alleles. Alleles in this group can be transcribed into an active messenger, despite the deletions of the promoter and the beginning of the coding sequence. The *FRI-Ler* allele however is transcribed to a much lower level than the alleles in the other groups (*FRI-fl* or *FRI-Col*), escaping detection by Northern blot as performed by Johansson et al (2000), but not by deep transcriptome sequencing in several accessions (Gan et al., 2011). The translation of the *FRI-Ler* transcript likely starts from the second in-frame methionine, resulting in a FRI protein lacking the first 42 amino acids, yet able to up regulate *FLC*, delay flowering time and confer vernalization sensitivity. In contrast, the *FRI-Col* allele is unable to confer any of these phenotypes, despite being expressed at a level comparable to the *FRI-fl* alleles present in most late flowering accessions.

Our experiment suggests that the reduced functionality of *FRI-Ler*, when compared to the wild type *FRI-Bil-7*, was mainly due to expression inhibition. We did detect small, non-significant but consistent differences between the *FRI-Ler* and *FRI-Bil-7* alleles when driven by the same promoter. This suggests that the deletion of the first 42 amino acids in FRI-Ler also affects its function, which is in agreement with what was proposed previously (Risk et al., 2010). In this work, Risk et al. demonstrated that transgenic Col-0 plants over-expressing a version of the FRI protein that lacks 118 residues from the N-terminal domain flowers earlier than a Col-0 line that contains an introgression of a functional *FRI* allele, although significantly later than Col-0. Since

the deletion in the *FRI-Ler* allele is much smaller than the one tested by Risk and collaborators, we can suggest that the *Ler* allele maintain at least some functionality.

In terms of the C-terminus end of the protein, the same study showed that, in contrast to what happens when the N-terminal deletion is present; the deletion of 151 residues from the C-terminus completely abolished the function of the FRI protein. Our work agrees with this result; showing all the nine tested *FRI* alleles with premature stop codons, including *FRI-Col*, are non-functional. Moreover, a study on the function of FRI has shown that the C-terminus 150 amino acids are essential for the interaction with its partner proteins *FLX*, *SUF4* and *FESI* (Choi et al., 2011); providing a mechanistic explanation for these results.

Many studies classify functionality of *FRI* alleles based only on the sequence prediction of a full-length protein (Atwell et al., 2010; Shindo et al., 2005; Stinchcombe et al., 2004). Only in a few cases, the functionality of the *FRI* alleles has been proved by genetic transformation or by crossing with accessions carrying non-functional *FRI* alleles (Gazzani et al., 2003; Johanson et al., 2000; Le Corre et al., 2002). In this thesis, I have cloned and expressed twelve *FRI* alleles encoding full-length proteins in Col-0 background and found that one of them, the *FRI-Wa-1* allele, is nonfunctional. Misclassification of this allele as functional can have negative consequences. For example, one report on flowering time in a F<sub>2</sub> population derived from a cross between Col and Wil-2, an accession carrying the *FRI-Wa-1* allele, proposed the existence of a dominant *FRI* suppressor closely linked to *FRI* in Wil-2 (Gazzani et al., 2003). The results from this thesis suggest that the locus affecting flowering time in this study is *FRI* itself but not a suppressor. We showed that the functional defect of the *FRI-Wa-1* allele is due to its coding polymorphisms. Further mutagenesis experiments confirmed that the single mutation from leucine to phenylalanine at residue 294 destabilizes the FRI protein and abolishes its function. During the course of my study, Hyun et al. (2016) published the crystal structure of the central domain of FRI from *Vitis vinifera* and pointed out that the leucine at position 294 is tightly surrounded by other hydrophobic amino acids; a substitution by a bulkier phenylalanine may undermine the structural stability of the entire protein. Our study provided experimental proof to this hypothesis.

*FRI* alleles have very high sequence diversity, especially in the first exon. This region has been proposed to be under strong selection to generate adaptively significant flowering time variation (Le Corre et al., 2002). The N-terminus of FRI is known to interact with FRIL1, which is



required for *FRI*-mediated up-regulation of *FLC* (Choi et al., 2011). It is possible that amino acid substitutions in this region influence protein function. In our experiments, most of the transgenic lines carrying functional *FRI* alleles contained mutations in this region and flowered late. Each transgenic line presents big variation in their flowering time, and there is no significant differences between lines. However, subtle differences in protein function caused by amino acid variation may not be detectable in our study, despite our efforts to minimize the variation caused by transgene insertion position and number by using large number of independent T<sub>1</sub> lines. Experiments with better control on the dosage of the transgene, or using introgression lines would be better suited to evaluate the effect of amino acids substitutions among full-length *FRI* alleles. Moreover, the effect of these amino acid substitutions on *FRI* protein function, if there is any, may be environment dependent.

Although *FRI* has a prominent role in the control of flowering time in Arabidopsis, not all variation observed among natural accessions can be explained by mutations in this gene. Previous studies have demonstrated that allelic variation at *FLC* itself accounts for variation in flowering time and adaptation (Irwin et al., 2016; Mendez-Vigo et al., 2011). As shown by (Caicedo et al., 2004), variation in flowering time may be associated with *FLC* only in the presence of a functional *FRI* allele. Consistent with this, we observed that accessions carrying the *FRI-Ler* or *FRI-fl* alleles, but not those carrying a *FRI-Col* allele, varied considerably in *FLC* expression and flowering time (Fig. 2-2B, 2-5B). Further variation of *FLC* levels can be due to additional factors such as mutations in genes of the autonomous floral promotion pathway, among others (Lee et al., 1994; Macknight et al., 1997; Simpson et al., 2003). Moreover, allelic variation at genes directly interacting with *FRI* could also have an impact in flowering time variation. For instance, it has been shown that *Ler* contains a non-functional *FRL1* allele, which is compensated by a functional *FRIGIDA-LIKE 2* (*FRL2*) allele. The latter restores *FRI*-mediated up regulation of *FLC* expression and, thus, a late flowering phenotype. In contrast, Col-0 carries a functional *FRL1* allele that is able to interact with *FRI*, but an impaired *FRL2* allele (Schlappi, 2006).

In summary, this work presents an updated and more precise functional classification of the *FRI* allelic series present in nature. The importance of variation in this gene for the life history of Arabidopsis makes the classification proposed here a powerful tool for future evolutionary studies in this species.

## 2.4 MATERIAL AND METHODS

### 2.4.1 Analysis of published datasets

Leaf number from Recombinant Inbred Lines (RILs) in the *Ler* x *Col* population (Clarke et al., 1995) was analyzed by grouping lines according to their genotypes at the closest molecular markers to *FRI* (m506, chr4 at 0.0cM) and *FLC* (g4560, chr5 at 17.3cM). Phenotypic and genotypic information for these individuals was kindly provided by Johan W. Van Ooijen and Caroline Dean.

Response to vernalization in multiple *Arabidopsis* accessions was obtained from published data (Lempe et al., 2005) and analyzed by subtracting days to flowering with vernalization from days to flowering without vernalization (grown in controlled environment rooms with 16 hour light). In this dataset, *FRI* alleles are genotyped as *FRI-Ler*, *FRI-Col*, functional alleles, novel alleles or none of the previous. For our analysis, we considered only those accessions classified as containing *FRI-fl*, *FRI-Ler* or *FRI-Col* alleles.

Expression for *FRI* and *FLC* was obtained from available repositories holding quantifications from northern hybridization for 167 *Arabidopsis* accessions grown in greenhouse under long day conditions (Atwell et al., 2010). Accessions in this dataset were classified according to their *FRI* alleles following the genotypes described in (Toomajian et al., 2006).

Fitness data was obtained from *Arabidopsis* accessions grown in four different field locations (Fournier-Level et al., 2011). These accessions were classified according to their *FRI* alleles as above.

*FRI* protein sequences were downloaded from GenBank. The species and sequence ID are listed in Table 2-1. The alignment was performed by Clustal Omega on the EMBL-EBI website.

**Table 2-1. The species and GenBank ID of *FRI* protein sequences used in the alignment.**

<i>Arabidopsis thaliana</i>	PODH90.1	<i>Boechera stricta</i>	AFJ66199.1
<i>Arabidopsis lyrata</i>	ABY51730.1	<i>Tarenaya hassleriana</i>	XP_010531851.1
<i>Capsella rubella</i>	AFJ66217.1	<i>Vitis vinifera</i>	XP_002283789.1
<i>Brassica rapa</i>	NP_001289004.1	<i>Medicago truncatula</i>	XP_003602718.2
<i>Camelina sativa</i>	XP_010426308.1	<i>Nelumbo nucifera</i>	XP_010250405.1
<i>Arabis alpina</i>	KFK36753.1		

### 2.4.2 Sequencing of *FRI* alleles

Twenty-three accessions (Alc-0, An-1, Bg-2, Bil-7, C24, Col-0, Cvi-0, Edi-0, *Ler*, Lip-0, NFA-10, NFA-8, Nok-3, Pro-0, Pu2-23, Ren-1, Sha, Spr-1-6, St-0, Ull-2-3, Van-0, Wa-1 and Zdr-6) were selected based on their *FRI* alleles (Shindo et al., 2005). Seeds of the accessions were obtained from the seed stock of Max Planck Institute for Plant Breeding Research. Genomic DNA was extracted from young leaves using the BioSprint 96 kit (Qiagen). A region of approximately 4.4k bp including the complete *FRI* gene and about 1.5k bp upstream and 0.6k bp downstream region was amplified in six overlapping fragments of 0.7k–1k bp using Phusion High-Fidelity DNA Polymerase (New England Biolabs). The location and sequence of the primers are listed in Table 2-2. Pooled PCR products from two independent reactions were purified using QIAquick PCR purification kit (Qiagen) and sequenced with same amplification primers via Sanger sequencing at the Max Planck Genome Centre Cologne. Individual sequencing results were assembled by SeqMan Pro (DNASTAR), aligned against the Col-0 reference sequence (TAIR10) and compared with the SNP information reported in Shindo *et al.* (2005).

**Table 2-2. Position and sequence of the primers used for cloning.**

Name	Position	Sequence	Name	Position	Sequence
Seq-F1	-1552	GGAAGCAAATGACCGTAAATC	Seq-R1	-582	AGGTGGGAAAGAAATGAAAACA
Seq-F2	-702	TGTTTCGTTGATGAGGGAGC	Seq-R2	+362	GATACATTGTTCCGCGGAG
Seq-F3	+283	AGAGTAACGGCGTTGTCCT	Seq-R3	+996	GATGTTGGTCGATGATGTCA
Seq-F4	+940	GTACCATATTCTGTTCTCACT	Seq-R4	+1673	TCGCAGGATCTAACTTGTGA
Seq-F5	+1602	GGCTACAAAGCAGCTAGCT	Seq-R5	+2305	ACATTCCTCCTATTTGGGGT
Seq-F6	+2249	CCAATCAAAGGTCTCCTCG	Seq-R6	+2946	TCTCAGTGCTGTATAACTACA

### 2.4.3 Cloning of *FRI* alleles

In order to clone the natural alleles of *FRI*, we amplified the genomic region containing *FRI* using Phusion High-Fidelity DNA Polymerase (New England Biolabs) by primer Seq-F1 and Seq-R6 (Table 2-2). The PCR product was digested with BamHI and EcoRI (New England Biolabs), whose restriction sites are at –1372 and +2691 respectively relative to the annotated start codon of *FRI* in TAIR10. The digested PCR product was separated by electrophoresis in a 1% agar gel and fragments with expected size were excised from the gel and purified by QIAquick Gel Extraction

Kit (Qiagen). The DNA fragments amplified from An-1, Col-0 and *Ler* were introduced into the binary vector pCAMBIA2300 between the BamHI and EcoRI restriction sites; while the fragments of all the other *FRI* alleles were introduced into pBinDsRed (kindly provided by Ed Cahoon, University of Nebraska), also making use of the BamHI and EcoRI restriction sites.

Promoter-swap constructs were generated using the MultiSite Gateway® Pro 2.0 system (Invitrogen) and genomic DNA from the accessions *Ler* (*FRI-Ler*) and Bil-7 (*FRI-Bil-7*). The *FRI-Ler* promoter region was cloned starting from position -1372, and ending at position +126, which is right before the first start codon downstream of the indel described in *FRI-Ler*. For *FRI-Bil-7*, the promoter region cloned ranged between positions -1372 and -1. Both promoter sequences were introduced into pDONR™ 221 P1-P5r by BP reaction. Similarly, the coding sequences together with 379 bp of their downstream sequence of both alleles (position +1 to position +2691 bp for *FRI-Bil-7* and position +127 to position +2691 for *FRI-Ler*) were cloned and introduced into pDONR™ 221 P5-P2. Finally, *FRI-Bil-7::FRI-Ler* and *FRI-Ler::FRI-Bil-7* constructs were generated by RL reaction that combined the entry vectors with the respective promoters and coding sequences into the binary vector pFAST1.

The *FRI-GFP* fusion protein was generated making use of binary vector pGWB404 (Nakagawa et al., 2007) that contains an in-frame GFP tag downstream of the gateway recombination site. *FRI* promoter and coding sequence up to the stop codon (-1372 to +2293) was amplified from the genomic DNA of Bil-7. A site directed mutagenesis from G to T at position +881, which results in a non-synonymous change from leucine to phenylalanine at amino acid 294, was introduced into the amplified *FRI-Bil-7* PCR fragment making use of overlapping primers (Ho et al., 1989). The wild type and mutated *FRI-Bil-7* PCR fragments were introduced into pGWB404 and named *FRI-Bil-7-GFP* and *FRI-L294F-GFP* respectively.

Inserts in all final constructs were sequenced and verified by comparison to the expected in silico constructs. Subsequently, all constructs were transformed into *Agrobacterium tumefaciens* strain GV3101 by electroporation. Agrobacteria transformed with pCAMBIA2300 or pBinDsRed were selected on solid YEP medium containing 50 mg/l kanamycin, 25 mg/l rifampicin and 50 mg/l gentamicin; whereas Agrobacteria transformed with pFAST1 or pGWB404 were selected on 100 mg/l spectinomycin, 25 mg/l rifampicin and 50 mg/l gentamicin. Positive colonies were confirmed by colony PCR and cultured in liquid YEP medium containing same antibiotics. In addition to

constructs mentioned above, an empty vector of pBinDsRed was transformed as negative control into Col-0 plants by the floral dip method (Clough and Bent, 1998).

**Table 2-3. Position and sequence of the primers used for cloning.**

Corresponding primers were used in cloning processes according to their positions as mentioned in the text. AttB sequence for the Gateway cloning is represented in red letters. The mutated nucleotides to introduce the L294P change are highlighted in green.

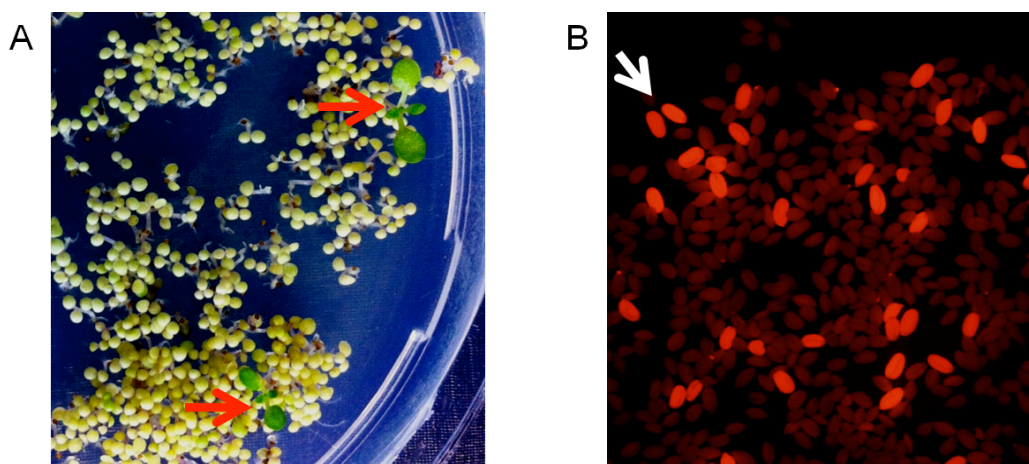
Name	Position	Sequence
pro-F	-1372	GGGGACAAGTTTGTACAAAAAGCAGGCTTAGATCCCAAATCTAGTGCCGCCG
pro-Ler-F	+126	GGGGACAACCTTTTGTATACAAAGTTGTACTTGTAAGCTCTGTTTCGACAATCTTC
pro-wt-F	-1	GGGGACAACCTTTTGTATACAAAGTTGTGAGATTGCGGCGAAGAAAAAGATTG
ter-Ler-F	+127	GGGGACAACCTTTGTATACAAAGTTGTAATGGACATTACGATCGGTCAATC
ter-wt-F	+1	GGGGACAACCTTTGTATACAAAGTTGTAATGTCCAATTATCCACCGACGGTG
ter-R	+2691	GGGGACCACTTTGTACAAGAAAGCTGGGTTCATCTTCCTCAAGAAAGATAAG
noSTP-R	+2293	GGGGACCACTTTGTACAAGAAAGCTGGGTTTTGGGGTCTAATGATGAGTTACTGC
mutL294-F	+896	GCTGGATTGATAAGGATGAGTGG
mutL294-R	+872	CCAATCATCCTTATAAATCCAGC

#### 2.4.4 Selection and phenotyping of transgenic plants

Transgenic plants carrying constructs in pCAMBIA (*FRI-Ler*, *FRI-Col* and *FRI-An-1*) or pGWB404 were selected on MS agar containing 50 mg/l kanamycin (Fig. 2-12A). Independent T<sub>1</sub> plants carrying a single T-DNA insertion were identified and propagated. Only T<sub>3</sub> lines that segregated at 3:1 ratio (resistant : susceptible) in the T<sub>2</sub> generation and line with homogeneous resistant phenotypes in the T<sub>3</sub> generation were used. All transgenic plants were grown in the greenhouse under long day conditions (16-hour light) with and without vernalization treatment for four weeks at 4°C.

Transgenic plants carrying constructs in pBinDsRed (*FRI* natural alleles) or pFAST1 (promoter swap between *FRI-Ler* and *FRI-Bil-7*) were selected by their red fluorescence under a fluorescence stereomicroscope (Leica MZ16F, Leica Microsystems) using green light of wavelength about 580 nm (Fig. 2-12B). T<sub>1</sub> plants carrying various *FRI* natural alleles were grown in the greenhouse under long day conditions for flowering time measurement. For the promoter swap experiment, T<sub>1</sub> lines carrying the *FRI-Bil-7*, *FRI-Bil-7::FRI-Ler* or *FRI-Ler::FRI-Bil-7* constructs were scored for flowering time under long day condition in an environmental chamber.

This experiment included plants from two independent T<sub>3</sub> lines carrying the *FRI-Ler::FRI-Ler* allele (L14 and L16 in Fig. 2-4B). Using T<sub>3</sub> lines allowed us to sow all the plants in the experiment directly on soil, as *FRI-Ler* T<sub>1</sub> lines do not have red fluorescent marker and would have required selection on MS agar containing kanamycin.



**Figure 2-12. Selection of T<sub>1</sub> transgenic plants.**

**(A)** Lines transformed with constructs in the pCambia2300 or pGWB404 vectors were selected on the MS agar medium containing 50 mg/l kanamycin. The red arrows point to the two transgenic plants that survived the antibiotic selection at 10 days after sowing. **(B)** Seeds from plants transformed with constructs in pBinDsRed or pFAST1 vectors were selected under fluorescence stereomicroscope. The red fluorescent seeds, such as the one pointed by the white arrow were considered transgenics.

Flowering time in all experiments was recorded as the number of days between seed sowing and flowering. The number of rosette leaves and cauline leaves were counted on the day when the first flower opened. The ratio between rosette and cauline leaves did not show any variation between the different genotypes in our experiment, so the total leaf number was used. We found a high correlation between total leaf number and flowering time in days, but differences between genotypes were larger when looking at leaf number. Consequently, only total leaf number is reported in this Thesis as a measurement of flowering time. Measurements of flowering time in the experiment with all T<sub>1</sub> transgenic plants containing different *FRI* alleles were terminated at end of

the 8<sup>th</sup> week. All plants not flowering at that time were assigned with the highest recorded total leaf number in the experiment.

### 2.4.5 Expression analysis using quantitative real-time PCR

For expression analysis, stratified seeds were sown on presoaked soil and grown in a growth chamber at long day condition for 10 days. Seedlings (leaves plus shoot) were harvested and immediately frozen in liquid nitrogen. Material from 8 to 10 seedlings was pooled for each one of the three biological replicates. Total RNA was extracted using Trizol (Ambion® TRIzol® RNA Isolation Reagent) or RNeasy Plant Mini Kit (QIAGEN). The DNA was removed from RNA by treatment with TURBO DNA-free™ Kit (Ambion). Complement DNA was synthesized from 5 µg RNA using Super Script® II Reverse Transcriptase in a 20 µl system (Invitrogen). Quantitative RT-PCR was performed on a CFX384 Touch™ Real-Time PCR Detection System using SYBR Green dye (iQ™ SYBR® Green Supermix, Biorad) using the primers in Table 2-4. The expression of *FRI* and *FLC* were normalized to the expression of *PP2A* and subject to further statistical analysis.

**Table 2-4. Primers used in the real-time PCR.**

Gene	qPCR-primer-F	qPCR-primer-R
<i>FRI</i>	TGCCTGATCGTGGTAAAGGGAAG	GCACCGGCAATCTCATTGGAAC
<i>FLC</i>	CCGAACATCATGTTGAAGCTTGTGAG	CGGAGATTTGTCCAGCAGGTG
<i>PP2A</i>	TAACGTGGCCAAAATGATGC	GTTCTCCACAACCGCTTGGT

### 2.4.6 Allele-specific expression analysis

Four F<sub>1</sub> hybrids generated by crossing *Ler* to four other accessions (Bor-4, Bur-0, Knox-10 and Sha) were grown in 12 h days in 12 h days in a controlled environmental chamber. Leaf material was collected at bolting stage. F<sub>1</sub> hybrids between *Ler* and Col-0 were grown in long days (16-hour light) in a controlled environmental chamber. Seedlings were collected 10 days after sowing. Genomic DNA was extracted from the collected material using DNeasy Plant mini kit (QIAGEN). cDNA from the five hybrids was generated using same method as described in the previous section.

PCR distinguishing *FRI-Ler* versus the four *FRI-wt* alleles targeted a nucleotide variant at position 235 relative to the start codon of *FRI* in TAIR10 that mutated from C (in Bor-4, Bur-0, Sha and Knox-10) to A (in *Ler*) (Table 2-5). To distinguish the *FRI-Ler* and *FRI-Col* alleles in the *Ler* x

Col F<sub>1</sub> hybrid, we targeted a variant at position 444 that changes from A in Col-0 to G in *Ler* (Table 2-5). Both cDNA and genomic DNA were used as template and the PCR was carried out with 50 cycles to eliminate all free primers.

Allele-specific mRNA abundances were measured using the PyrosequencerAB (Biotage AB, Qiagen) following manufacturer's instructions for sample preparation and pyrosequencing reactions. Vacuum sample preparation was performed using 15 µL of PCR product mixed with 5 µL of Streptavidin Sepharose beads (GE Healthcare), 40 µL of PyroMark Binding Buffer (Qiagen) and 20 µL of LiChroSolv water (Merck). Pyrosequencing was performed in a PyroMark Q96 Plate Low (Qiagen) in each well containing 1 µL of sequencing primer (10 µM) and 40 µL of Pyromark Annealing Buffer (Qiagen) using PyroMark Gold Q96 Reagents (Qiagen). Allele-specific expression for each SNP was estimated using the pyrosequencing software (PSQ 96MA 2.1.1, Biotage AB) based on the peak height for each allele at the SNP. A peak correction factor of 0.86 was used for incorporation of dATP αS, as recommended by the manufacturer.

**Table 2-5. Primers used in the allele specific analysis.**

	<b>Chr4: 269,257</b>	<b>Chr4: 269,469</b>
amplification primer-F	Biotin-TCAGTTGCACTGGAAACATTCA	Biotin-ATTGTACCGGAGACGTCGAATAA
amplification primer-R	GCGTTTTTCGATTGACTCGATGT	GGCCAATTTCAAAGCTGAAG
pyrosequencing primer	TGACTCGATGTGCTTCT	CTTGCTACACATCAACTC

### 2.4.7 Microscopic imaging

Seeds were sterilized and sown on MS agarose medium and grown vertically under long day condition for 7 days. At this time, whole seedlings were dipped in a solution of propidium iodide (10 mg/ml) for one minute and rinsed in water. The root tips were cut off and mounted for visualization under confocal microscope (ZEISS LSM800).

### 2.4.8 Western blot assay

Seedlings from the imaging experiment detailed above were also used for western blot analysis. For this, seedlings were pooled and pulverized in liquid nitrogen. Total protein was extracted using PEB protein extraction buffer (Agrisera Antibodies). The protein solution was adjusted to equal concentration and separated on a NuPAGE™ 4-12% Bis-Tris Protein Gel (Thermo Fisher Scientific). Each sample was loaded twice on the gel, resulting in two identical gels



after separation. One gel was subjected to Coomassie staining and the other was used for blotting. Anti-GFP and secondary antibodies were used (Roche) according to their protocol. Detection was carried out using the Amersham ECL Prime Western Blotting Detection Reagent (GE healthcare) following their protocol and band intensity was analyzed with Image J software.

## Chapter 3

---

# Identification of *VIP HOMOLOG2 (VIH2)* as a novel flowering time gene underlying a flowering time QTL

### Summary

A new flowering locus at the top of chromosome 3 (named QTL3) was identified through QTL analysis of flowering time in a recombinant inbred line (RIL) population derived from a cross between *Ler* and An-1.

We studied the presence of QTL3 in various segregating populations and found it in three populations that had *Ler* as a common parent (*Ler* x An-1, *Ler* x Sha and *Ler* x Eri). We obtained Near Isogenic Lines (Gross et al.) from all three populations and confirmed the presence of the QTL. The *Ler* allele at QTL3 delayed flowering in all cases. Fine mapping of QTL3 with the *Ler* x Eri NILs narrowed down the locus to a single gene: *VIP HOMOLOG 2 (VIH2)*, AT3G01310). *VIH2* encodes a kinase that catalyzes the synthesis of inositol pyrophosphates (InsP<sub>7</sub> and InsP<sub>8</sub>) and is well conserved in all eukaryotic organisms.

We measured InsP<sub>8</sub> in the NILs and found higher abundance associated to the *VIH2-Ler* allele. Mutant lines for *vih2* are early flowering, further proving the role of *VIH2* in flowering time regulation, and implicating that *VIH2-Ler* is a gain-of-function mutant allele. In agreement with this, we found that the *VIH2-Ler*-NIL is also more responsive to jasmonate signaling, a pathway that linked InsP<sub>8</sub> signaling to flowering time.

### **3.1 INTRODUCTION**

In the previous chapter of this thesis we identified *FRI* as the candidate gene for a flowering time QTL located in chromosome 4 in a recombinant inbred line (RIL) population derived from a cross between the accessions An-1 and *Ler* (Tisne et al., 2010) (Fig. 2-1). In this section, we focus on the identification of an additional flowering time locus located on top of chromosome 3 (QTL3) in the same population. As with the QTL for *FRI*, QTL3 was detected in plants grown under 12-hour photoperiods both in control conditions and under water deficit, where half of the soil water content was used.

We investigated the presence of QTL3 in other QTL analysis of flowering time in *Arabidopsis*. We found evidence of QTL3 in various populations involving *Ler* as a parent. For example, QTL3 was found in three RIL populations derived from crosses between *Ler* and the Asian accessions Kas-2, Kond and Sha (el-Lithy et al., 2006; El-Lithy et al., 2004). In the *Ler* x Kas-2 and *Ler* x Kond populations, QTL3 was present when plants were grown in long day photoperiods (16-hour light / 8-hour dark), while in the *Ler* x Sha population QTL3 was only found under 12-hour light photoperiods, but not under 16-hour light condition. Among populations that combine *Ler* and other European accessions, QTL3 was reported in a RIL population derived from *Ler* x Eri, but only under specific stress conditions (low temperature and high intensity light in long day photoperiods) (Ilk et al., 2015). In all these RIL populations, the *Ler* allele at QTL3 delays flowering. We also found evidence of QTL3 in two RIL populations not involving *Ler*: the Col x Sha and Bay-0 x Sha populations. In both cases, Sha alleles at QTL3 result in early flowering of RILs, although in the Bay-0 x Sha RILs the QTL is only present when RILs were grown in high-density condition but not in control conditions (Botto and Coluccio, 2007; Simon et al., 2008). In summary, QTL3 is a robust locus that influences flowering time in multiple populations of *Arabidopsis*. Inspection of the region of QTL3 did not reveal any obvious candidate gene for this locus.

In the following chapter, we confirm QTL3 in three populations: *Ler* x Sha, *Ler* x An-1 and Eri x *Ler*. Fine mapping using Near Isogenic Lines (Gross et al.) from the *Ler* x Eri population lead us to a single candidate gene, *VIP HOMOLOG2* (*VIH2*). *VIH2* contains a conserved ATP-grasp kinase domain and catalyzes the synthesis of inositol pyrophosphates (InsP<sub>7</sub> and InsP<sub>8</sub>), which represent a group of versatile signaling molecules. Sequence as well as biochemical analysis of the

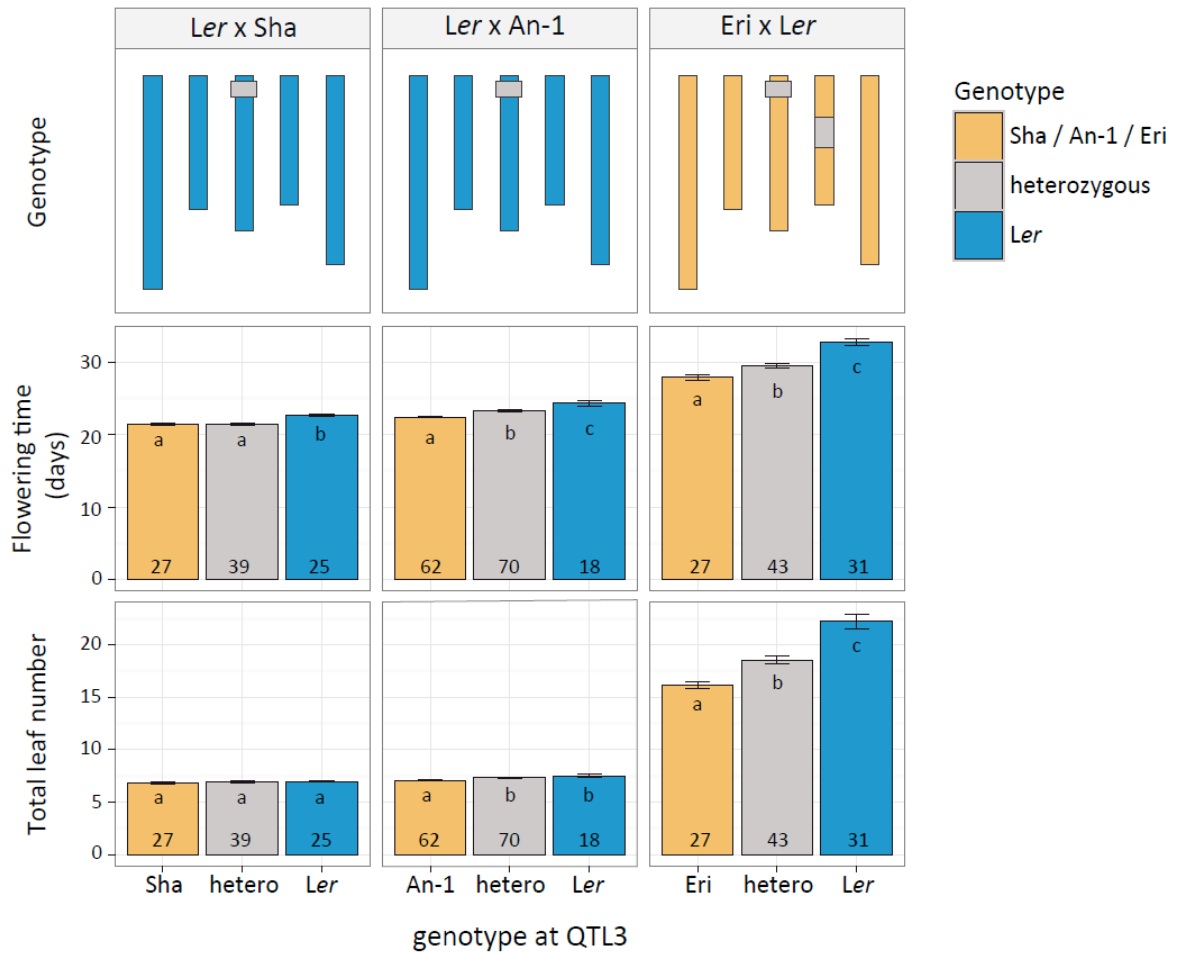
natural alleles involved suggests functional variation caused by coding sequence polymorphisms. Characterization of lines segregating for *VIH2-Eri* or *VIH2-Ler* alleles, as well as *vih2* mutants and over-expressing lines confirm the identity of *VIH2* as the gene underlying QTL3. This is the first time that *VIH2* is implicated in the control of flowering time and that natural variation is described for this gene.

## 3.2 RESULTS

### 3.2.1 Confirmation of QTL3 in multiple populations

As mentioned in the introduction, QTL3 has been reported in multiple populations where the *Ler* accession was involved as a parent. My colleague Inga Schmarlenbach developed a Near-Isogenic Line (NIL), which carries the introgression on the top of chromosome 3 from An-1 accession into the *Ler* genetic background. Inga also obtained two other NILs with introgression at the QTL3 region. The NIL with Sha allele of QTL3 introgressed into *Ler* genome is kindly provided by Dr Carlos Alonso Blanco. The other NIL carries *Ler* allele of QTL3, which is introgressed into the Eri accession genomic background (developed by Nadine Ilk). The NILs were crossed to their respective genetic background parental accession to establish lines segregating at QTL3 that would allow us to fine map and confirm the effect of QTL3 on flowering time (Fig. 3-1). Among this set of lines, two NILs contained clean *Ler* backgrounds except for small introgressions of An-1 or Sha alleles on top of chromosome 3. The third NIL consisted of an Eri background with *Ler* introgressions on the top of chromosome 3 and middle of chromosome 4.

Plants that descend from all three heterozygous lines were grown in the greenhouse under long day condition, and number of days from sowing and total leaf number were scored at the day of the opening of the first flower. The *Ler* allele at QTL3 delays flowering time in all lines (Fig. 3-1). In the early flowering *Ler* x Sha population, the flowering time difference is only significant for the number of days between sowing and opening of the first flower, but not in the number of total leaves at flower opening. The difference in flowering time is most prominent in the Eri NILs, likely because of their general late flowering phenotype induced by the Eri background. It is worth noticing that a small region in chromosome 4 is segregating in these NILs in addition to the region of QTL3, but the flowering time differences only co-segregated with the genotype at QTL3.

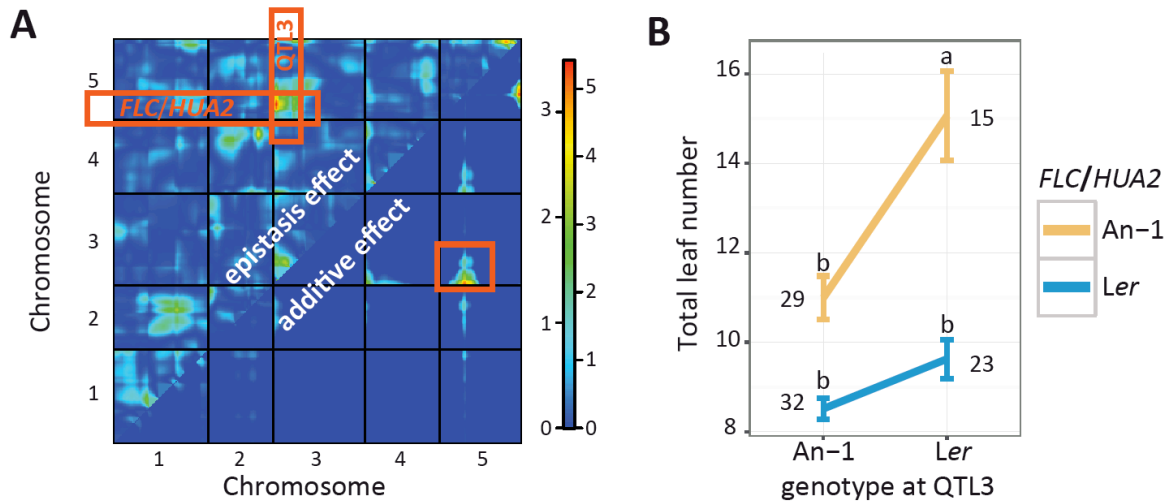


**Figure 3-1. Confirmation of QTL3 in three NILs.**

Heterozygous NILs were derived from populations between *Ler* and *Sha*, *An-1* or *Eri*. The lines segregating on top of chromosome 3 were grown in a greenhouse under long day conditions. Each plant was genotyped at QTL3 and scored for flowering time, both in number of days between sowing and the opening of first flower and in the number of total leaves on the day when the first flowering opened. Error bars represent the standard error of the mean. The number at the bottom of each bar indicates the number of individual plants analyzed per line. Letters in each bar represent significance groups as determined by the Tukey HSD test.

The original QTL analysis on *Ler* x *An-1* RIL population shows that QTL3 partially interacts with another QTL at the top of chromosome 5 (Fig. 3-2A). This region contains two closely linked flowering time genes known to affect flowering time in this population, *FLC* and *HUA2* (Doyle et al., 2005; Michaels et al., 1999). Both of these genes are functional in the *An-1* accession, whereas

in *Ler*, *FLC* shows a reduced function due to an insertion of a 1.2kb transposable element in its first intron (Michaels et al., 2003) and *HUA2* is non-functional because of a pre-mature stop codon (Doyle et al., 2005). The reduced function of these genes in *Ler* was confirmed in the *Ler* x An-1 RIL population, where lines carrying *Ler* alleles at these loci exhibit early flowering (Fig. 3-2B). Interestingly, the flowering time difference caused by QTL3 is only significant in the presence of functional An-1 allele of *FLC* and *HUA2* (two-way ANOVA,  $p < 0.001$  interaction between QTL3 and multiple markers on top of chromosome 5). In the NIL with an Eri genetic background, both *FLC* and *HUA2* are functional, greatly enhancing the effect of QTL3. In the subsequent study, we used the *Ler* x Eri NILs to fine-map the causal gene underlying QTL3.



**Figure 3-2. Interaction between QTL3 and top of chromosome 5 in the *Ler* x An-1 RIL population.**

**(A)** Two-dimensional QTL analysis on the flowering time in the *Ler* x An-1 RIL population. LOD scores reflecting epistatic interactions are presented in the upper left triangle; LOD scores reflecting additive effect without epistasis are presented in the lower right triangle. In the color scale for LOD scores, the numbers on the left and right side correspond to the scale of upper left triangle and lower right triangle respectively. **(B)** Flowering time of RILs grouped by the genotype on top of chromosome 3 and 5. All plants were grown under 12-hour light condition in the PHENOPSIS automated phenotyping platform and measured at the time of the opening of the first flower (Tisne et al., 2010). Error bars represent the standard error of the mean. The number on the side indicates the number of RILs in each genotype; only RILs with same genotype at *FLC* and *HUA2* were selected for the analysis. Letters on the top represent significance groups as determined by the Tukey HSD test.

### 3.2.2 Characterization of the *Ler* x *Eri* NILs.

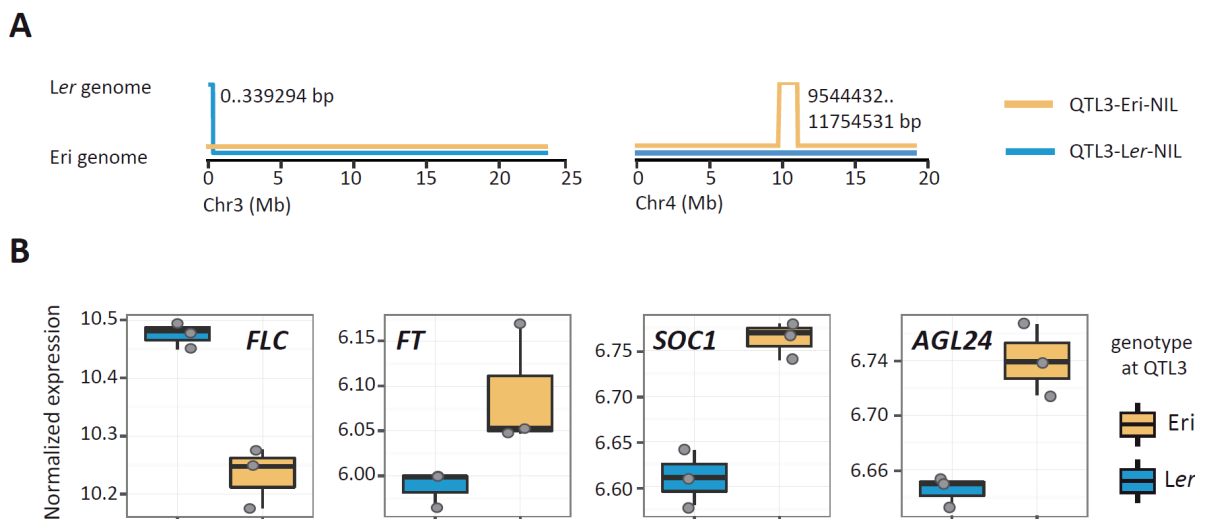
We studied in detail the *Ler* x *Eri* NILs because of the larger phenotypic variation observed in these lines in the previous section (Fig. 3-1). For this, we performed an RNA-seq experiment using two homozygous NILs that differed in their genotype at the QTL3 region; this would allow us to genotype and phenotype for expression levels. The *Eri*-NIL and *Ler*-NIL (named according to their genotype at QTL3) were grown under 16-hour light photoperiods and 10-day-old seedlings were harvested for RNA-seq analysis. We sequenced three biological replicates for each NIL and obtained an average of 22 million reads per sample (range 18 to 26 million) and used TopHat2 to align 88% percent of the reads to the TAIR10 Arabidopsis reference.

We first confirmed the genotype of the NILs with the RNA-seq data. For this, we run a standard pipeline in the lab that uses GATK to call SNPs and short insertion/deletions (indels) in exonic regions. Marker analysis in the NILs showed that indeed both NILs are homozygous throughout their genomes and are mostly identical, except for the two expected regions in chromosomes 3 and 4 (Fig. 3-3A). The *Eri*-NIL carries homozygous *Eri* allele on top of chromosome 3, but *Ler* allele in the middle of chromosome 4. The *Ler*-NIL carries the opposite genotypes at both regions. This analysis restricted the location of QTL3 between 0 and 339,294 bp on chromosome 3, containing 126 protein-coding genes; none of them are known to control flowering time. Analysis of the effect of SNPs and indels in the proteins located in the QTL3 region did not reveal any candidate with obvious deleterious mutations or modification. We then analyzed the sequence of An-1, *Eri*, Sha, Kond and Kas-2 accessions (re-sequencing data available from the 1001 genome project). All the four accessions are parents in the RIL populations, in which *Ler* allele of QTL3 delays flowering (as described in the introduction). We looked for polymorphisms that differentiated *Ler* from all other accessions. Although no deleterious or splicing site mutations were found with these characteristics in the 126 genes in the region, most genes contained non-synonymous mutations that could give rise to functional differences.

We next used the RNA-seq results to calculate differential expression and alternative splicing between the two NILs. Again, we found no gene with evidences of differential expression or alternatively splicing between NILs among the 126 candidates in the region of QTL3. We therefore concluded that one or more of the abundant non-synonymous polymorphisms located in the region is likely the cause of the flowering time differences between the NILs.



Outside the candidate region in chromosome3, the RNA-seq analysis identified 980 differentially expressed genes, including various well known flowering time genes (Fig. 3-3B). The flowering repressor *FLC* is expressed at higher level in the *Ler*-NIL; this observation supports the genetic interaction between QTL3 and *FLC* previously observed in the QTL analysis performed on the *Ler* x An-1 RILs (Fig. 3-2). The flowering promoter genes *FT*, *SOC1* and *AGL24*, whose expression are repressed by *FLC*, show higher expression level in the Eri-NIL, in agreement with the earlier flowering time phenotype found in this line (Fig. 3-1).

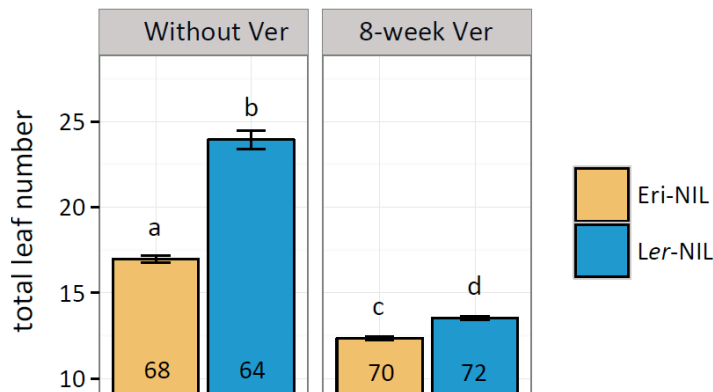


**Figure 3-3. RNA-seq analysis of Eri-NIL and Ler-NIL.**

**(A)** Both Eri-NIL and Ler-NIL have identical homozygous genomes, differing only in two regions on chromosome 3 and 4. The genotype and the position of the differed regions are labeled in the schematic representation of chromosome 3 and 4. **(B)** Normalized expression of four flowering time genes differentially expressed between the two NILs. Material used for RNA-seq analysis was taken from aerial parts of 10-day-old seedlings grown in long day conditions. Three biological replicates (pools of several plants) were used for each genotype.

We investigated the genetic interaction between QTL3 and *FLC*, whose expression is silenced by vernalization. For this, we characterized flowering time of the NILs grown in long day conditions with and without 8-week-vernalization. As shown in Fig. 3-4, the *Ler* allele of QTL3 delays flowering in both vernalized and non-vernalized plants. As expected, vernalization promotes

flowering in both NILs and greatly reduced the differences in flowering time between the lines; although a small but significant difference persisted. This experiment suggests that the regulation of flowering time by QTL3 is only partially dependent on *FLC*, and there is also a *FLC* independent mechanism. In terms of photoperiod, NILs differing at a larger region at QTL3 retain their flowering time differences when grown under 8-hour-light short day condition (N. Ilk, personal communication), suggesting that QTL3 is not involved in the photoperiod pathway.



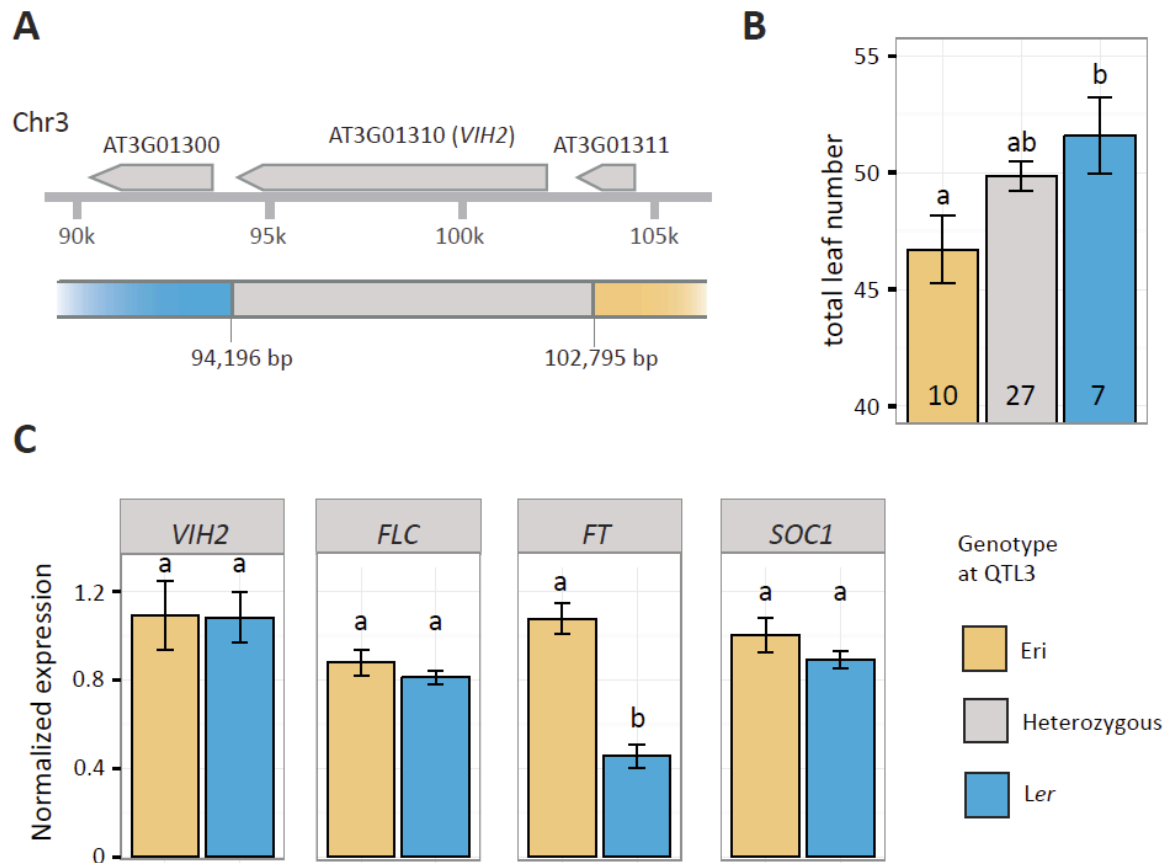
**Figure 3-4. Effect of vernalization on the flowering time of Eri and Ler-NILs.**

All plants were grown in greenhouse under long day conditions. For vernalization, 10-day-old plants were transferred to 4 °C and short days and released back to the greenhouse after 8 weeks. Error bars represent the standard error of the mean. The number in each bar indicates the number of individual plants per line analyzed. Letters on top of each bar represent significance groups as determined by the Tukey HSD test.

### 3.2.3 Identification of *VIH2* as causal gene underlying QTL3

We used the heterozygous Eri NILs described above to fine map QTL3. We designed markers on the borders of the introgression and searched for seedlings that presented a recombination event inside the QTL3 region. Recombinant lines were scored for flowering time, and heterozygous lines that segregated for flowering time were used for further fine mapping steps. After screening about 3500 plants in three sequential fine mapping steps, we obtained a line with a small heterozygous region between 94.2 kbp and 102.8 kbp on chromosome 3 that still segregated for flowering time. The borders of the heterozygous region are located at the promoter and

terminator regions of the *VIH2* (AT3G01310) gene (Fig. 3-5A). Flowering time of the descendants of this line segregate according to the genotype at *VIH2* (Fig. 3-5B), unequivocally pointing at *VIH2* as the gene underlying QTL3. We named the two NILs differing to each other only the *VIH2* gene as the *VIH2-Eri-NIL* and *VIH2-Ler-NIL* according to their genotype at *VIH2*.



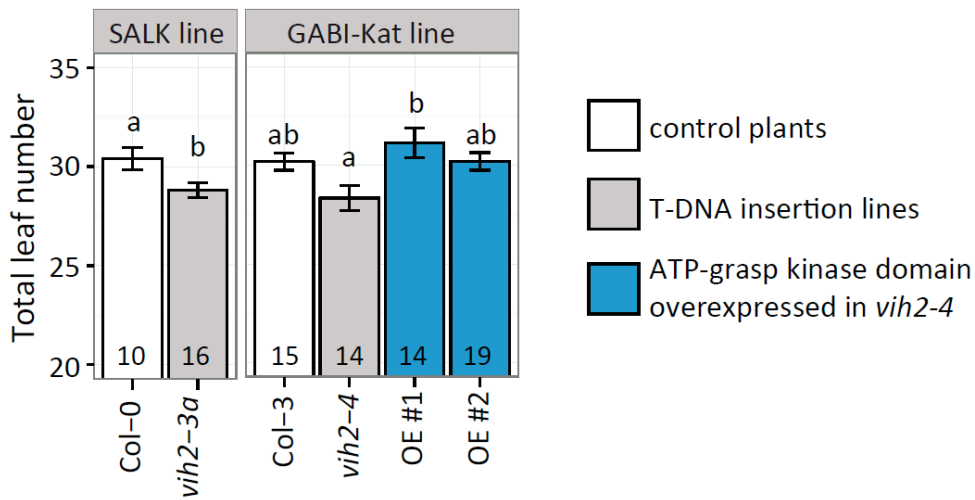
**Figure 3-5. Identification of *VIH2* as causal gene underlying QTL3**

**(A)** The position of the heterozygous region in the backcrossed *VIH2*-NIL relative to the location of genes on chromosome 3. **(B)** Flowering time of the progenies from the heterozygous individual as shown in (A) when grown under long day condition in a greenhouse. **(C)** Expression of *VIH2* and flowering time genes in the homozygous lines as shown in (B), expression was normalized to the expression of *PP2a*. Material was taken at ZT10 from aerial parts of 10-day-old seedlings grown in an environmental chamber under long day condition. Three biological replicates each constituted by a pool of multiple plants were used for qPCR for each genotype. Error bars represent the standard error of the mean. The number in each bar indicates the number of individual plants per line analyzed. Letters in each bar represent significance groups as determined by the Tukey HSD test.

RNA-seq analysis of the NILs segregating for an introgression spanning 126 genes evidenced differences in expression in a number of flowering time genes but not in any of the genes located in introgression (Fig. 3-3 and Section 3.2.2). We aimed to confirm these results by qRT-PCR in the *VIH2*-NILs (Fig. 3-5C). As seen before in the RNAseq experiment, the expression of *VIH2* is similar in both NILs. To our surprise, we did not observe differences in *FLC* expression as we did before between the NILs differing in 126 genes, indicating that these differences could be due to a different gene in the QTL3 region. We tested the expression level of two flowering time integrator genes *FT* and *SOC1*; and found significantly higher expression in *FT* in the earlier flowering *VIH2-Eri*-NIL, but no difference in *SOC1* expression. This is the first expression analysis performed on the *VIH2*-NILs; the plant materials were harvested for analysis on *FLC* expression at about ZT10, before the peak expression of *FT* at ZT16 (Valverde et al., 2004); however, we could still observe the significant difference in *FT* expression. Further experiments will be conducted to confirm these observations.

We studied the regulatory role of *VIH2* on flowering time by analyzing two *vih2* mutant lines that have been previously characterized by Laha et al. (2015). The SALK T-DNA line *vih2-3a* and the GABI-Kat T-DNA line *vih2-4* were grown together with their wild type controls, Col-0 and Col-3 respectively, in a greenhouse under 12-hour light photoperiods. Both mutant lines flowered earlier than their wild-type controls, though the difference is only significant in the *vih2-3a* line (Fig. 3-6).

*VIH2* encode a kinase catalyze the synthesis of inositol pyrophosphate; it is comprised of the N-terminal ATP-grasp domain and the C-terminal. The expression of ATP-grasp domain alone is enough to restore the inositol pyrophosphate synthesis in the *vih2-4* mutant (unpublished data from Dr Gabriel Schaaf's group). The constitutive over-expression of the ATP-grasp kinase domain of *VIH2* restored, and even delayed flowering time of the *vih2-4* mutant compared to the wild type. The results suggest that *VIH2* function as a negative regulator of flowering time in Arabidopsis. If this is true, we can infer that the observed flowering time differences between the *VIH2-Eri* and *VIH2-Ler* NILs could be due to increased activity of the *VIH2-Ler* allele.



**Figure 3-6. *VIH2* delays flowering in the Col background.**

*VIH2* mutant lines flowered earlier than their respective wild types. Transgenic lines carrying *35S::VIH2-ATP-grasp-kinase-domain* in the *vih2-4* mutant flowered later than the mutant background. All plants were grown under 12-hour photoperiods in the greenhouse. The number in each bar indicates the number of individual plants per line analyzed. Letters above each bar represent significance groups as determined by the Tukey HSD test.

### 3.2.4 Protein sequence and functional variation between *VIH2-Eri* and *VIH2-Ler* alleles

Because the expression of *VIH2* is the same in both *VIH2*-NILs, the difference in flowering time is likely caused by mutations in the protein sequence. Compared to the reference *VIH2* sequence in Col-0 (UniProtKB ID: F4J8C6.1), *VIH2-Eri* carries a conservative amino acid change at position 484 where the negatively charged aspartic acid residue in Col-0 is replaced with another negative charged residue glutamic acid (D484E). Because of the similar charge of both residues this mutation is unlikely to alter the protein function. In contrast, *VIH2-Ler* carries a dramatic amino acid change at position 213. In this case, the polar threonine residue in Col-0 or Eri is replaced by a hydrophobic isoleucine residue (T213I). In addition, this mutation occurs in the ATP-grasp kinase domain, which, on its own, is sufficient to delay flowering time in transgenic assays (Fig. 3-6).

We studied the T213I amino acid change present in the *VIH2-Ler* allele among *VIH2* homologs. *VIH2* is part of the VIP family of proteins, which are highly conserved in all eukaryotic

organisms (Desai et al., 2014). We aligned the sequence of VIP homologs from various species whose function has been experimentally confirmed. These sequences include the two VIP homologs from *Arabidopsis*, one from the green alga *Chlamydomonas reinhardtii*; three animal VIP homologs from human, mouse and rat; the budding yeast and fission yeast VIP homologs as well as two additional sequences from fungi (Couso et al., 2016; Menniti et al., 1993; Mulugu et al., 2007; Pohlmann et al., 2014; Stephens et al., 1993). The threonine residue at position 213 that is mutated in *VIH2-Ler* is highly conserved both in plants and animals. In all fungi VIP homologs, except budding yeast, position 213 is valine (Fig. 3-7). The chemical property of the isoleucine residue in *VIH2-Ler* is similar to valine, so the *VIH2-Ler* protein is most likely functional. Given that *VIH2-Ler* allele delays flowering in comparison with *VIH2-Eri*, and that the null *VIH2* mutants present early flowering, we hypothesize that the amino acid mutation in *VIH2-Ler* increases the activity of the enzyme.

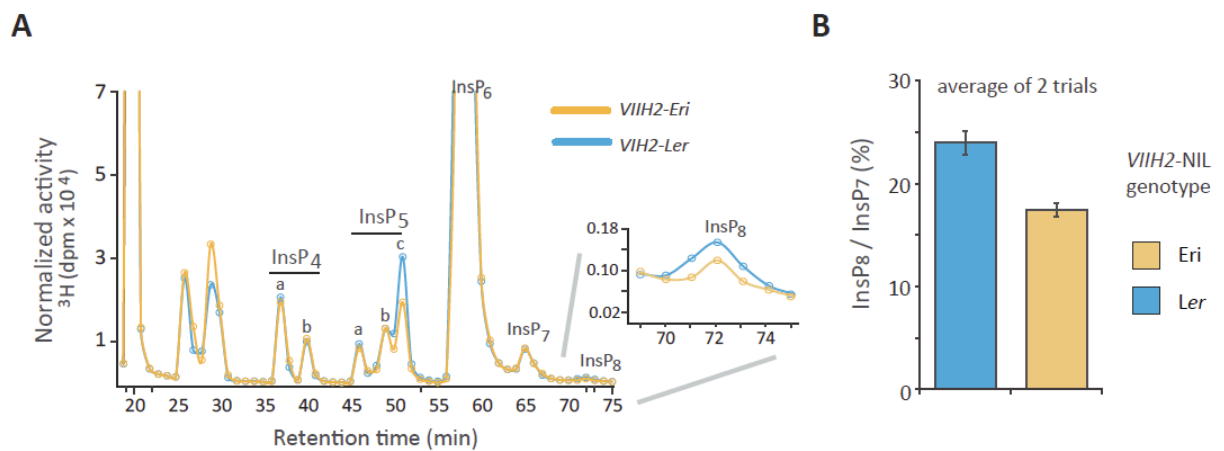
AtVIH2	EFHPDV-R---RVRREGSYIYEEFMA <b>T</b> GG-TDVKVYTVGPEYAHAEARKSPVVDGVVMRN	245	plant
AtVIH1	EFHPDV-R---RVRREGSYIYEEFMP <b>T</b> GG-TDVKVYTVGPEYAHAEARKSPVVDGVVMRN	244	
<i>Chlamydomonas</i>	KYDSEHNG---CVRRDGSFIYEEFLP <b>T</b> GG-TDVKVYTVGPRYAHAEARKSPVVDGKVLRS	257	green alga
mouse_VIP1	VYSPES-----SVRKTGSYIYEEFMP <b>T</b> DG-TDVKVYTVGPDYAHAEARKSPALDGKVERD	285	animals
rat_VIP1	VYSPES-----SVRKTGSYIYEEFMP <b>T</b> DG-TDVKVYTVGPDYAHAEARKSPALDGKVERD	287	
human_VIP1	VYSPES-----SVRKTGSYIYEEFMP <b>T</b> DG-TDVKVYTVGPDYAHAEARKSPALDGKVERD	285	
mouse_VIP2	VYSPES-----NVRKTGSYIYEEFMP <b>T</b> DG-TDVKVYTVGPDYAHAEARKSPALDGKVERD	280	
human_VIP2	VYSPES-----NVRKTGSYIYEEFMP <b>T</b> DG-TDVKVYTVGPDYAHAEARKSPALDGKVERD	274	fungi
budding yeast	EFDPVLVHPT----EGSYIYEQFMD <b>T</b> DNFEDVKAYTIGENFCHAEATRKSPVVDGIVRRN	440	
fission yeast	DYDPDLCAPT----EGSFIYEEFMNVDAEDVKVYTVGPHYSHAETRKSPVVDGIVRRN	286	
<i>Ustilago maydis</i>	EYDPNLVEPT----DGSYIYEEFMDVDNAEDIKVTIGPHFVHAETRKSPVVDGVVKRN	258	
<i>Aspergillus nidulans</i>	EYDPDLRTPRSILEDGSSYIYEQFLRVDNAEDVKAYTVGPDFCHAEATRKSPVVDGLVRRN	383	
:			.*.***:.*: ... *.**.*: : **.****:.* * *

**Figure 3-7. Sequence alignment of VIP homolog proteins.**

Protein sequences of the VIP proteins whose function have been experimentally proven were aligned with Clustal Omega. Proteins are grouped according to the organisms they belong to. The threonine residue mutated in *VIH2-Ler* allele is labeled in red.

We studied differential functionality of the *Ler* and *Eri* alleles of *VIH2* in collaboration with the Schaaf lab at the University of Tübingen. The *VIH2* protein catalyzes the synthesis from InsP<sub>6</sub> to Ins P<sub>7</sub> and InsP<sub>7</sub> to Ins P<sub>8</sub>. The product to substrate ratio of an enzyme can reflect its catalytic function. Given the high abundance of InsP<sub>6</sub> in plants, the InsP<sub>7</sub> / InsP<sub>6</sub> ratio is relatively stable; whereas both InsP<sub>7</sub> and InsP<sub>8</sub> have very low abundance in plants; and their ratio is a good indication of *VIH2* enzymatic activity (Desai et al., 2014). It has been shown that *Arabidopsis* lines

lacking the *VIH2* protein show an elevated level of  $\text{InsP}_7$  and a very low level of  $\text{InsP}_8$  (Laha et al., 2015), thus a reduced  $\text{InsP}_8 / \text{InsP}_7$  ratio. We measured the abundance of  $\text{InsP}_7$  and  $\text{InsP}_8$  in seedlings from the *VIH2*-NILs. In two consecutive trials, the *VIH2-Ler*-NIL consistently shows higher  $\text{InsP}_8$  to  $\text{InsP}_7$  ratio than its sibling containing the *Eri* allele (Fig. 3-8). This result agrees with our hypothesis that the *Ler* allele of *VIH2* encodes for a more active protein than the *Eri* allele. A more direct assay on enzyme function will be carried out in yeast in the future.



**Figure 3-8. *Ler* encodes an enzymatically more active allele of *VIH2*.**

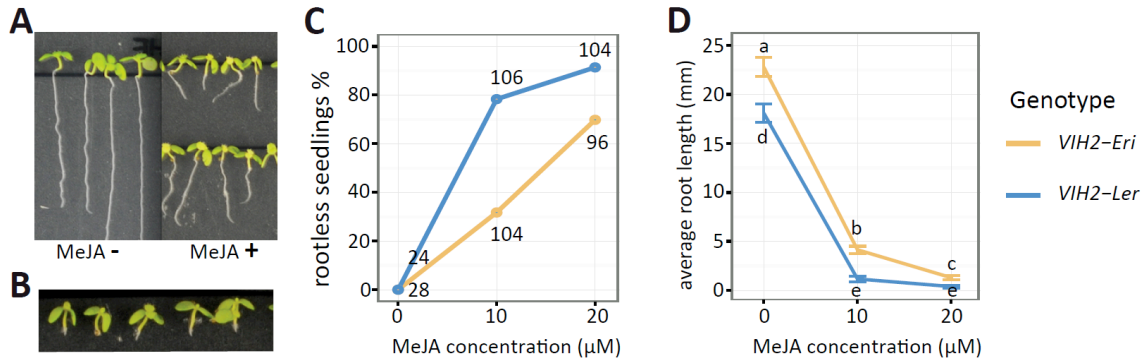
**(A)** Normalized HPLC profiles of inositol phosphate species of 3-week-old [ $^3\text{H}$ ] inositol-labeled *VIH2*-NIL seedlings. Extracts were resolved by Partisphere SAX HPLC and fractions collected each minute for subsequent determination of radioactivity. The inset figure shows the enlarged view at the peak of  $\text{InsP}_8$ . The experiment was repeated twice with comparable results. **(B)** The average  $\text{InsP}_8 / \text{InsP}_7$  ratio from the two experiments.

### 3.2.5 *VIH2-Ler* NIL is more sensitive to jasmonate

It has been proposed that *VIH2* is involved in jasmonate-dependent defense (Laha et al., 2015). *VIH2* respond to jasmonate signal and increase the production of  $\text{InsP}_7$  and  $\text{InsP}_8$ , which bind to the ASK1-COI1-JAZ jasmonate receptor complex to facilitate downstream signaling.

We investigated whether the *VIH2*-NILs differing only in their *VIH2* alleles show variation in their responses to jasmonate. A classical method to test jasmonate responses is through measurement of primary root growth, that is known to be inhibited by Jasmonate (Staswick et al.,

1992). We grew the *VIH2*-NILs on agar medium containing methyl jasmonate (MeJA) and compared root lengths of 10-day-old seedlings. MeJA inhibited root growth in seedlings from both the *VIH2-Ler* and *VIH2-Eri*-NILs (Fig. 3-9A). On the medium containing 10 or 20  $\mu$ M MeJA, a considerable number of seedlings from both *VIH2*-NILs exhibited complete inhibition on root growth (Fig. 3-9), but the percentage of the rootless seedlings was significantly higher in the NIL carrying the *VIH2-Ler* allele (Fig. 3-9C). At 20  $\mu$ M MeJA, almost all the *VIH2-Ler*-NIL seedlings were rootless, whereas only 70% *VIH2-Eri*-NIL seedlings were rootless. We calculated the average root length of all the seedlings based on their genotype and the MeJA concentration, the rootless seedlings were considered with root length at 0 mm. The root lengths were significantly shortened by MeJA in the medium, and the shortening is more extensive in the *VIH2-Ler*-NIL seedlings due to the high percentage of the rootless seedlings (Fig. 3-9D). These observations indicate that the *VIH2-Ler* allele confers higher responsiveness to jasmonate, which could be due to the higher  $\text{InsP}_8$  content observed before in this line (Fig. 3-8). It would be interesting to repeat the experiment with lower MeJA concentrations that do not inhibit the root growth completely; it is also our interest to carry out the root length assay using the *vih2* mutants and over-expressing lines.



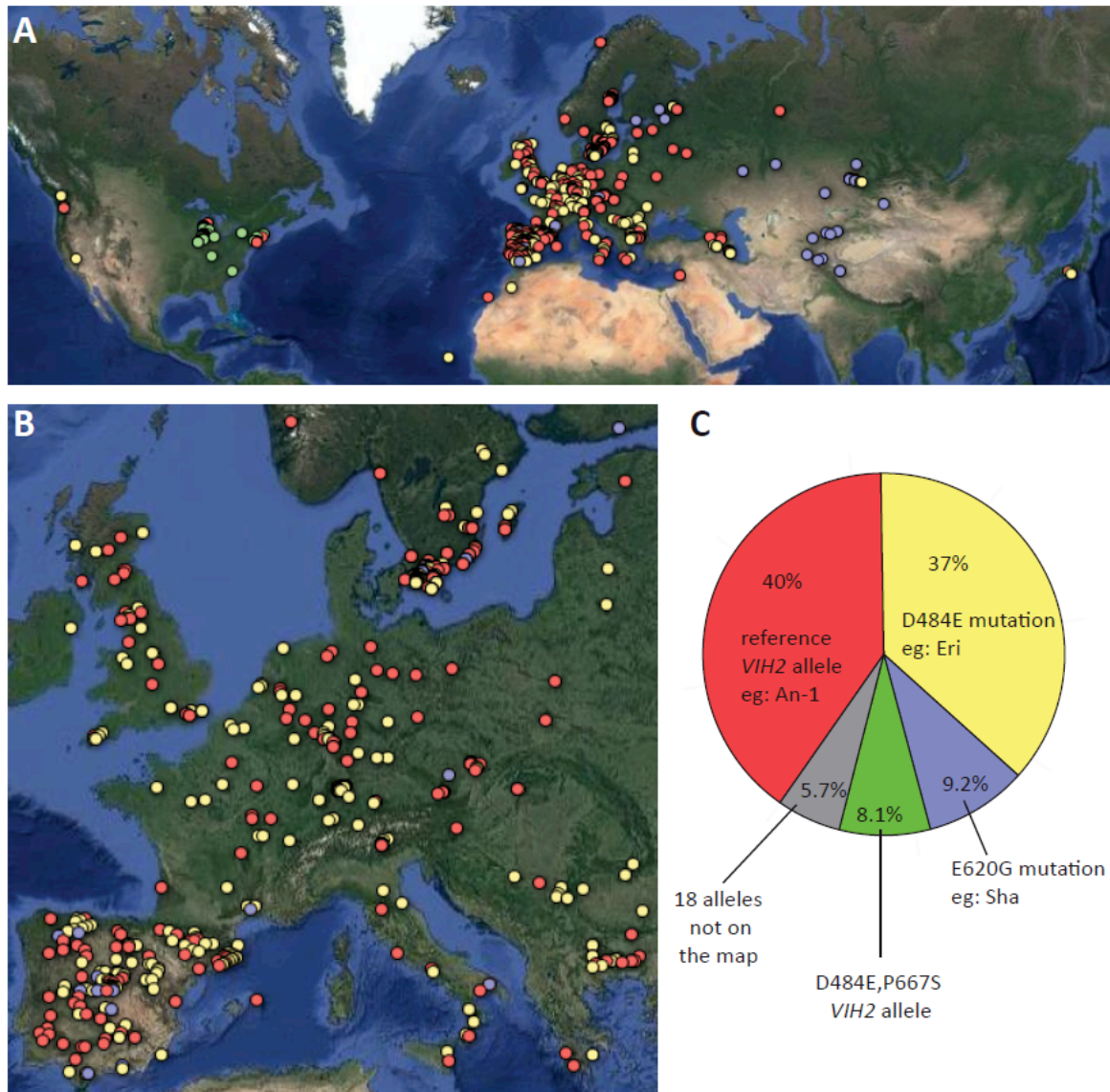
**Figure 3-9. Root growth of *VIH2*-NILs on MeJA containing agar medium.**

**(A and B)** Seedlings growing on MeJA containing medium exhibit root partial growth inhibition (A) or complete root growth inhibition (B). Seedlings from both *VIH2-Eri* or *VIH2-Ler*-NILs show these phenotypes in the presence of 10  $\mu$ M or 20  $\mu$ M MeJA. **(C)** Percentage of seedlings with complete inhibition of root growth among the *VIH2*-NILs. The numbers represent number of seedlings examined in the experiment in each genotype. **(D)** The average root length of seedlings with MeJA treatment. The rootless seedlings were counted as with root length of 0 mm. The letters at each point represent significance groups as determined by the Tukey HSD test.



### **3.2.6 *VIH2* allele variation and distribution among *Arabidopsis* accessions**

We studied the distribution of *VIH2* alleles in nature. Interestingly, the T213I amino acid mutation likely responsible for the increased enzymatic activity of the *VIH2-Ler* allele is not present in any other sequenced accession, including the closely related *Ler-1*. As it could be expected from its conservation across kingdoms, the sequence of the *VIH2* protein is highly conserved in *Arabidopsis*. Two alleles differing in a single amino acid are present in 77% of all 1135 accessions sequenced in the 1001 genome project (The 1001 Genomes Consortium, 2016) (Fig. 3-10). This is a mutation from aspartic acid at position 484 present in the reference accession Col-0 to glutamic acid, as observed in accessions such as Eri. These two *VIH2* alleles are distributed throughout Europe (Fig. 3-10). Sequence comparison with predicted *VIH* proteins from other Brassicaceae revealed that the ancestral allele of *VIH* carries glutamic acid at position 484, as in the Eri allele. The third most abundant allele can be found in accession from Europe and Asia and is characterized by a replacement of the negatively charged glutamic acid residue at position 620 by a side-chain-free glycine residue (E620G). This allele is the most prevalent among Asian accessions, including Sha, Kas-2 and Kond. The fourth most common *VIH2* allele derives from the ancestral allele present in Eri and carries an additional mutation from proline to serine at position 667 (D484E, P667S). This allele is restricted to accessions from eastern United States. Seventeen other *VIH2* alleles were found in the 1135 accessions, but they are present at very low frequency. Together with *VIH2-Ler*, these 18 alleles occur only in 5% of all the accessions. So far, the only allele for which we have evidences of functional differences is the *Ler* allele described in this thesis.



**Figure 3-10. Geographic distribution of the common *VIH2* alleles.**

**(A)** Collection site for accessions carrying the four most common *VIH2* alleles in nature. Each dot represents an accession and the color represents the *VIH2* allele they carry. **(B)** Zoom-in map of Europe, where the *Arabidopsis* accessions are most densely collected. **(C)** Allele frequency of the most common *VIH2* alleles among the 1135 *Arabidopsis* accessions sequenced in the 1001 genome project. Colors correspond to the color on the map.

### 3.3 DISCUSSION

#### 3.3.1 Identification of *VIH2* as the causal gene of flowering time QTL3

In this part of the thesis, we characterized a QTL for flowering time in chromosome 3 found in the *Ler* x An-1 RIL population (Tisne et al., 2010). Examination of published flowering time QTL analyses showed the existence of this QTL in different RIL populations grown under various conditions, indicating that the effect of the underlying locus on flowering time is very robust. Interestingly, at least 5 of the populations that presented QTL3 had *Ler* as a common parent, suggesting that the *Ler* allele was rather unique.

We confirmed QTL3 using three sets of NILs that contained introgressions at the QTL3 region alternating *Ler* alleles with An-1, Sha or Eri alleles. In all of these lines, the presence of *Ler* alleles was associated to later flowering time. Because of the larger effect of QTL3 in the *Ler* x Eri NILs, we characterized in detail and conducted fine mapping in these lines.

In the QTL analyses conducted in the *Ler* x An-1 and *Ler* x Eri RIL populations, QTL3 interacts genetically with the *FLC* locus in chromosome 5, showing both epistasis and additive effect (Ilk et al., 2015; Tisne et al., 2010). This interaction could be confirmed using RNAseq data from the *Ler* x Eri NIL lines, where the expression of *FLC* was higher in the late flowering NIL carrying the *Ler* allele for QTL3. Interestingly, the late flowering phenotype of this NIL persists after the expression of *FLC* is silenced by vernalization, allowing us to conclude that QTL3 is involved in flowering time regulation both in a *FLC*-dependent and -independent manner.

After three rounds of fine mapping we were able to obtain a heterozygous NIL segregating at a single gene in the region of QTL3: *VIH2*. The descendants of this NIL presented co-segregation between their flowering time and the genotype at *VIH2*; thus pinpointing at *VIH2* as the causal gene underlying QTL3. The *vih2* mutant lines in Col-0 background also support its role in flowering time regulation. However, it is worth noticing that the *VIH2*-NILs, segregating for a single gene, flowered much later (Fig. 3-5B) than the NILs initially used for fine mapping, that contained 126 genes (Fig. 3-1) (45~55 total leaves compared to 15~25 leaves at flowering). In addition, the expression of *FLC* is not different between the *VIH2*-NILs carrying Eri or *Ler* alleles of *VIH2*. It is thus possible that additional gene(s) controlling flowering time through regulation of *FLC* are in the region of QTL3; whereas *VIH2* seems to regulate the expression level of *FT* independently of *FLC*.

### 3.3.2 Function of *VIH2* and its product: inositol pyrophosphates in plants

*VIH2*, the gene underlying flowering time QTL3, is a homolog of the yeast VIP1 kinase that catalyzes the synthesis of inositol pyrophosphates (Desai et al., 2014; Laha et al., 2015).

Inositol pyrophosphates are sub-group of inositol phosphate molecules (InsP) based on *myo*-inositol, which is a six-carbon-ring with a hydroxyl group on each carbon (cyclohexane). Any of the six hydroxyl groups in *myo*-inositol can be phosphorylated in a combinatorial manner, resulting in a family of structurally diverse stereoisomers. When all the six hydroxyl groups are phosphorylated (InsP<sub>6</sub>), further phosphorylation leads to inositol pyrophosphates. Multiple inositol phosphate molecules have been identified in plants. The most abundant form, InsP<sub>6</sub>, accumulates highly in seeds; comprising more than 1% of the dry weight of the seed and occupying two-thirds of the total amount of seed phosphorus (Raboy et al., 1984). For this reason, it was assumed that phosphorus storage is the key role of inositol phosphate in plant (Raboy and Dickinson, 1987). Inositol phosphate gained more attention after the discovery of InsP<sub>6</sub> bound to auxin receptor complexes and InsP<sub>5</sub> in jasmonate receptor complexes (Sheard et al., 2010; Tan et al., 2007). The inositol pyrophosphate molecules were only identified in plants recently (Desai et al., 2014), and overall, they represent less than 2% of the total inositol phosphates in plants.

In *Arabidopsis*, there are two highly-conserved VIH paralogs, named as *VIH1* and *VIH2*. The expression of *VIH1* is restricted to pollen, whereas *VIH2* is ubiquitously expressed, and especially high in rosette leaves. For this reason, it is thought that *VIH2* is responsible for the conversion of InsP<sub>6</sub> into InsP<sub>7</sub> and subsequently to InsP<sub>8</sub> in most tissues (Laha et al., 2015). Study on the function of inositol pyrophosphates in plants has just started. To date, the only reported function of inositol pyrophosphates in *Arabidopsis* is linked to the COI1-JAZ jasmonate receptor complex. COI1 is an F-box protein in the SKP1-CUL1-F-box protein (SCF) ubiquitin E3 ligase complex. In the presence of the bioactive JA, COI1 binds to the jasmonate ZIM-domain (JAZ) transcriptional repressors; triggers the polyubiquitylation and the subsequent degradation of JAZ repressors by the 26S proteasome, which relieves the transcription of JA responsive genes (Chini et al., 2007; Katsir et al., 2008). Previously, InsP<sub>5</sub> has been shown to bind to the COI1-JAZ complex (Sheard et al., 2010). Laha et al. (2015) suggested that the more anionic InsP<sub>8</sub> is the preferred binding partner of COI1-JAZ complex in planta, because COI1 and JAZ proteins form together a binding pocket that fits InsP<sub>8</sub>, which would act as a chemical glue to stabilize the complex in the presence of JA. The

*vih2* mutant is less responsive to elevated JA; the mutant lacks InsP<sub>8</sub> but contains high levels of InsP<sub>7</sub> and, further proving that InsP<sub>8</sub>, and not InsP<sub>7</sub>, facilitates the binding between COI1 and JAZ. In summary, production of InsP<sub>8</sub> by *VIH2* facilitates jasmonate dependent defense against wounding and pathogen attacks.

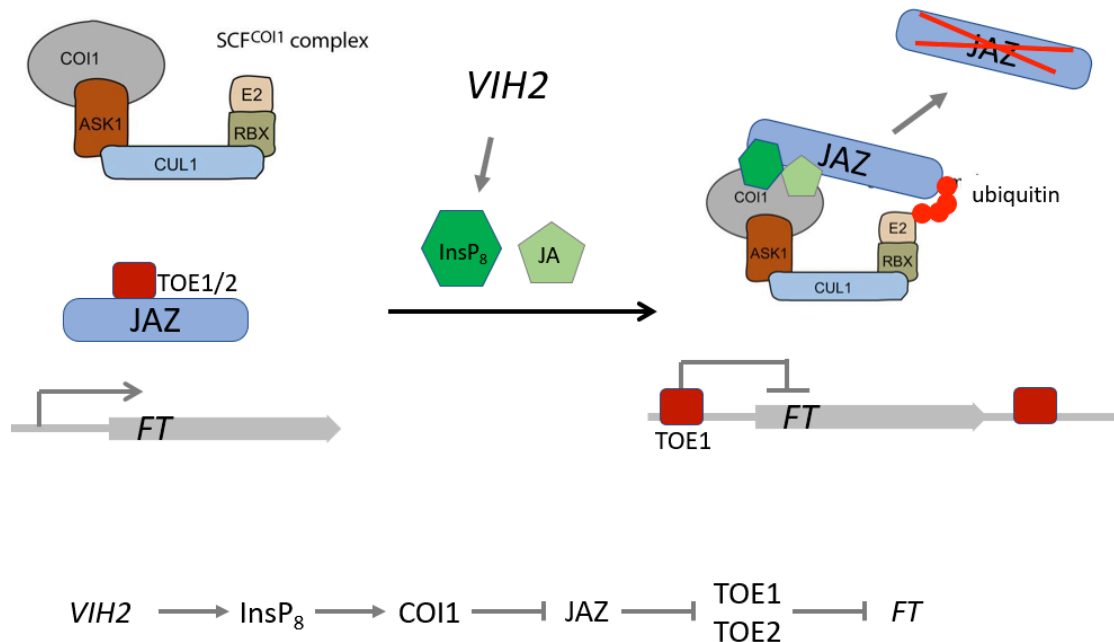
In our research, we found that the *Ler* allele of *VIH2* enhanced enzymatic activity and enabled higher InsP<sub>8</sub> accumulation in planta. Preliminary results suggest that the NIL with *VIH2-Ler* is more sensitive to external JA treatment, implicating that the elevated InsP<sub>8</sub> level indeed facilitates the binding of COI1 to JAZ proteins, which leads to more JAZ degradation and stronger JA response. It is therefore likely that the *VIH2-Ler* allele grants the NIL with stronger resistance against herbivore and pathogen attack, but further research will be needed to determine if this is the case.

### **3.3.3 The role of *VIH2* in flowering time regulation.**

One hypothesis about the molecular mechanisms used by *VIH2* to control flowering could be through its involvement in JA signaling. The core component of the JA receptor complex, COI1, has been shown to play a role in flowering time regulation. The *coil-1* mutant is insensitive to JA and presents early flowering (Song et al., 2013). Zhai et al. (2015) reported the link between COI1 and the flowering time integrator *FT*. The JAZ proteins in the JA signaling pathway can physically bind the proteins of flowering repressors: *TARGET OF EARLY ACTIVATION TAGGED 1 (TOE1)* and *TOE2*, this binding inactivates TOE1 and TOE2 proteins. In the presence of jasmonate, the JAZ proteins undergo COI1-mediated degradation. Consequently, the TOE1 and TOE2 proteins are released from the JAZ sequester and bind to the promoter and downstream region of *FT* genetic locus; inhibiting the transcription of *FT*. The overall effect is that the activation of JA signaling pathway delays the flowering of plant by repressing *FT* transcription.

Based on the current knowledge, we propose a hypothetical molecular mechanism to explain the flowering time differences observed in our study (Fig 3-11). Our results suggest an increased enzymatic activity of *VIH2-Ler* allele that leads to elevated accumulation of InsP<sub>8</sub>, which in turn leads to more efficient COI1-dependent degradation of JAZ repressors. The lower abundance of JAZ proteins in *VIH2-Ler*-NIL causes higher activity of JAZ target proteins, among which are the

*FT* repressors TOE1 and TOE2. In the *VIH2-Ler-NIL*, TOE1 and TOE2 proteins would be free of inhibitory binding of JAZ and would physically associate with the promoter and downstream region of *FT* locus and inhibit its transcription. This inhibition of *FT* in the *VIH2-Ler-NIL* would explain its late flowering compared to the *VIH2-Eri-NIL*. This proposed pathway mostly relies on post-translational regulation that we could not verify experimentally with our current plant materials. However, we did observe low expression of *FT* in the *VIH2-Ler-NIL*, whose transcriptional regulation is the last step in our model. This mechanism depends on functional JA biosynthesis and COI1, so the genetic interaction between *VIH2* and COI1 or JA biosynthesis would need to be tested.



**Figure 3-11. Proposed molecular mechanism of *VIH2* regulation on flowering.**

### 3.3.4 Natural variation of *VIH2* among *Arabidopsis* accessions

Our study suggests that T213I mutation present in the *Ler* allele of *VIH2* may be the cause of its enhanced enzymatic activity. The *VIH2-Ler* allele is unique among all 1135 *Arabidopsis* accessions sequenced in the 1001 genome project (The 1001 Genomes Consortium, 2016). Interestingly, the sister accession to *Ler*, *Ler-1*, does not contain this mutation. It has been proposed

that the *erecta* mutation in *Ler* is the result of a gamma ray induced mutagenesis (Koncz et al., 1992), thus it is possible that the T213I mutation was caused by the same mutagenesis treatment.

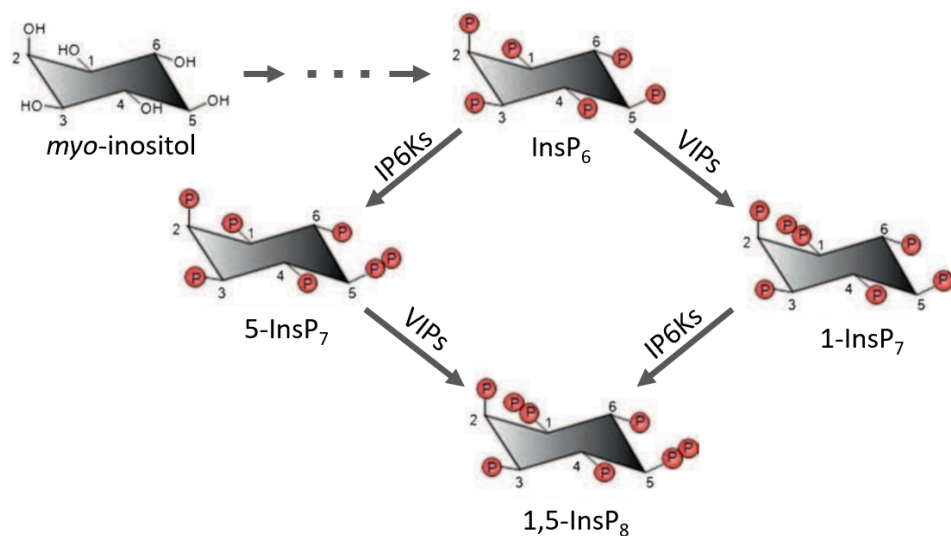
The structure and function of VIH2 proteins are conserved throughout eukaryotic organisms. As expected, the sequences of the VIH2 proteins are highly conserved among Arabidopsis accessions. About 80% of the accessions in the 1001 genome project carry two major alleles, which differ in only one amino acid at residue 484, aspartic acid or glutamic acid, represented by Col and Eri respectively. Compared to the predicted VIH proteins in other Brassicas species, the ancestral allele of VIH carries glutamic acid at position 484, as in the allele found in the Eri accession. The accessions carrying these two common alleles are found throughout Europe and usually mingled together, so the mutation probably occurred very early in time, before the spreading of Arabidopsis. Given the conservative amino acid change and the equal geographic distribution of the two alleles, it is most likely that the resulting proteins have identical enzymatic functions. The third most common allele is the one present in the Sha accession; it carries one mutation compared to the *VIH2-Col* allele: the negatively charged glutamic acid residue at position 620 is replaced by side-change-free glycine residue. The *VIH2-Sha* allele is found in accessions located both in Europe and Asia. It is worth noticing that the *VIH2-Sha* allele absolutely dominates in the accessions found in central Asia. One flowering time QTL analysis using Col x Sha RIL population identified QTL3 as a flowering locus (Simon et al., 2008). If *VIH2* is indeed the causal gene, this study implicates the *VIH2-Sha* allele functionally differs from *VIH2-Col* and leads us to wonder whether the altered enzymatic activity confer fitness or if its predominance in Asia is merely due to founder effect. Further studies are needed to distinguish both possibilities. The forth most common allele of *VIH2* carries a mutation from proline to serine at position 667, when compared to the ancestral *VIH2-Eri* allele. This allele is found exclusively in accessions from eastern United States, and all the accessions belong to the genetic group from Germany. The mutation may have occurred among the Arabidopsis settled in America very early, where Arabidopsis spread very quickly, so the unique geographic distribution is more likely due to a founder effect than to positive selection. In the future, we will test the enzymatic function of the common *VIH2* alleles by expressing them in yeast. Differential enzymatic function may open further discussion on the role of VIH2 in adaptation.

### 3.3.5 Pleotropic effect of inositol pyrophosphates in yeast and animals

The role of inositol phosphates as signaling molecules has long been studied in yeast and animals. The best understood among them is the  $\text{Ca}^{2+}$  release factor 1,4,5- $\text{InsP}_3$  (Streb et al., 1983). Inositol pyrophosphates have also been discovered earlier in the slime mold (Stephens et al., 1993) and rat cell lines (Menniti et al., 1993) than in plants. In these systems, inositol pyrophosphates are important metabolic regulators, maintaining energy homeostasis. In yeast, inositol pyrophosphates participate in regulation of cellular ATP concentration (Szigyarto et al., 2011); at the whole organism level in mammals, they involve in insulin signaling to regulate glucose uptake (Chakraborty et al., 2010). Other studies have shown that inositol pyrophosphates play a role in phosphate nutrient regulation in yeast and animal cells (Lee et al., 2007; Saiardi et al., 1999). In addition, they also involve in innate immune response in mammals (Prasad et al., 2011; Pulloor et al., 2014). In summary, as reviewed by Thota and Bhandari (2015), inositol pyrophosphates can be considered important signaling molecules that regulate a good number of biological processes. In my Thesis, we find a role for these molecules in flowering time regulation in Arabidopsis.

The biosynthesis of inositol pyrophosphates has been studied in detail in budding yeast and mammalian cells (Mulugu et al., 2007; Saiardi et al., 1999; Saiardi et al., 2001). Two conserved groups of enzymes catalyze the addition of phosphate groups at specific positions in  $\text{InsP}_6$ : the  $\text{InsP}_6$  kinases (IP6Ks) at position 5 and the VIPs at position 1. Both enzymes take  $\text{InsP}_6$  as substrate to produce 1- or 5- $\text{InsP}_7$ , and they subsequently phosphorylate each other's products to generate 1,5- $\text{InsP}_8$ , as shown in Fig.3-12. The *VIH* proteins in Arabidopsis are *VIP* homologs, which show high structural and functional conservation with the yeast *VIP1*. The expression of Arabidopsis *VIH1* or *VIH2* in the yeast strain with *vip1* mutation restores the *VIP* function, catalyzing phosphorylation at position 1, but not 5, in the inositol ring (Laha et al., 2015). It is worth noticing that, in both yeast and mammals, synthesis of  $\text{InsP}_8$  requires both *VIP* and *IP6K* enzymes. A null mutation in either *VIP* or *IP6K* resulted in higher accumulation of  $\text{InsP}_7$ , but complete lack of  $\text{InsP}_8$ , similar to the  $\text{InsP}$  profile we observed in the *vih2* mutant. However, no plant *IP6K* homolog has been identified based on sequence homology (Desai et al., 2014). The fact that *vih2* mutant have higher  $\text{InsP}_7$ , but no  $\text{InsP}_8$ , implicate an unidentified enzyme carrying out the function of *IP6K* in plants.





**Figure 3-12. Synthesis of inositol pyrophosphate from  $\text{InsP}_6$  in yeast and animal.**

The inositol molecule is presented in chair conformation. The position of each carbon is numbered. Monophosphate groups are represented by red spheres. Two connected red spheres indicate position of the pyrophosphate groups. (Figure adapted from Thota and Bhandari (2015))

### 3.3.6 Outlook for pleiotropy of *VIH2* in Arabidopsis

Given the extensive pleiotropic effect of inositol pyrophosphates in yeast and animals, we expected a similar feature in plants. A QTL analysis on abundance of  $\text{InsP}_6$  and inorganic phosphate ( $\text{Pi}$ ) using a RIL population derived from a cross between *Ler* and *Cvi* found a QTL co-localizing with QTL3 (Bentsink et al., 2003). In that study, the *Ler* allele is responsible for lower  $\text{InsP}_6$  and  $\text{Pi}$  in both leaf and seeds (Bentsink et al., 2003). A later QTL analysis on seed  $\text{InsP}_6$  abundance using *Ler* x *Eri* and *Ler* x *Kond* RIL populations revealed a locus at same position (Ghandilyan et al., 2009). In both RILs, the *Ler* allele at QTL3 is responsible for lower  $\text{InsP}_6$  in seeds. In all studies, these QTLs present major effects, explaining 20% ~ 60% of the respective phenotypic variance. *Cvi* carries the *VIH2* allele as *Eri*, so *Ler* carries a different allele of *VIH2* in all these phosphate related QTL analysis. Based on our work, we suggest *VIH2* is a possible candidate for these QTLs. In the *Ler* x *Cvi* population mentioned above, QTL3 was found to affect  $\text{Pi}$  in addition to  $\text{InsP}_6$ . This could be attributed to the role of inositol pyrophosphate in  $\text{Pi}$  sensing and regulation, which has been described before. In yeast, the  $\text{Pi}$ -responsive (PHO) pathway component Pho4 binds to the intragenic region of the yeast IP6K homolog and lower the  $\text{InsP}_7$  level

in response to low Pi conditions (Nishizawa et al., 2008). In frog oocytes, the expression of IP6K leads to increased Pi uptake (Norbis et al., 1997). In plants, inositol pyrophosphates have been suggested to participate in Pi homeostasis, but this has not been directly proven (Kuo et al., 2014).

Interestingly, *VIH2* was pinpointed as a candidate gene for a QTL controlling vitamin E accumulation in seeds present in the *Ler* x *Cvi* RIL population (Gilliland et al., 2006). In this case, the researchers fine mapped the locus to genetic region containing only *VIH2* and its upstream gene AT3G01300, which is an unidentified protein kinase. Currently, there is no link established between inositol phosphate signaling and vitamin E biosynthesis; but this is mostly because research on inositol phosphate is so far conducted in yeast or mammalian systems, which do not produce vitamin E. It will be interesting to test differences in the level of vitamin E in the seeds of *VIH2*-NILs or the *vih2* mutants.

In brief summary, we identified *VIH2* as the causal gene underlying a flowering time QTL and proposed a molecular pathway to explain the role of *VIH2* as a flowering time repressor. Study in yeast and animals have shown that the products of *VIH2*, inositol pyrophosphates, are signaling molecules with diverse pleotropic effects. Research on inositol pyrophosphate in plants has just started, and it has been proven to participate in JA-dependent defense. Here we have proven its role in flowering time regulation and raise the possibility of its implication in other processes of biological importance.

### 3.4 Materials and Methods

#### 3.4.1 Plant materials and phenotyping of flowering time

Three NILs used in the confirmation of QTL3 were derived from crosses between *Ler* x An-1, *Ler* x Sha and *Ler* x Eri. The *Ler* x An1 NILs were generated in our lab by Dr. Inga Schmalenbach and the latter two were kindly provided by Dr Maarten Koornneef and Dr Carlos Alonso-Blanco respectively. The *vih2* T-DNA insertion lines used in this study were *vih2-3a* (SAIL\_165\_F12) and *vih2-4* (GK-080A07); as characterized by Laha et al. (2015). The *VIH2* over-expression lines are transgenic lines carrying *35S::VIH2-ATP-grasp kinase domain::eYFP* construct in the *vih2-4* mutant background, kindly provided by our collaborator Dr Gabriel Schaaf.

In all experiments, plants were grown in a greenhouse under photoperiods as specified in the results session. For the vernalization experiment, seedlings were grown in greenhouse for 10 days under a 16-hour light photoperiod and then transferred to a climatic chamber set up at 4 °C and 8-hour photoperiods. After 8 weeks in the vernalization chamber, seedlings were put back to the 16-hour photoperiod greenhouse. Control plants (non-vernalized) in the vernalization experiment were sown in the same greenhouse 10 days before the release of the vernalized plants from their treatment. In all experiments, flowering time was measured as total leaf number.

#### 3.4.2 Genotyping at QTL3

NILs with introgression at QTL3 were backcrossed to their respective genetic background. In the segregating F<sub>2</sub> population, flowering time of individual plants was recorded and the region of QTL3 genotyped with marker nga172.

An F<sub>2</sub> population derived from the *Ler* x Eri NIL was used for fine mapping of QTL3. the recombinant line containing an introgression of only the *VIH2* gene (*VIH2*-NIL), the *VIH2* genomic region was amplified with over-lapping primers and Sanger sequenced to confirm the size of the introgression. Primers for fine mapping and sequencing are listed at the end of this chapter.

#### 3.4.3 Expression analysis using RNA-seq and quantitative real-time PCR

For expression analysis, 10-day-old seedlings were grown under 16-hour light days in an environmental chamber. Seedlings (leaves plus shoot) were harvested and immediately frozen in liquid nitrogen. Material from 8 to 10 seedlings was pooled for each one of the three biological

replicates. Total RNA was extracted using RNeasy Plant Mini Kit (QIAGEN). DNA was removed from each sample with the TURBO DNA-free™ Kit (Ambion). For RNA-seq analysis, libraries were prepared according to the Illumina TruSeq RNA protocol and sequenced on the Illumina HiSeq2000 platform (Illumina, Inc.) at the Genome Center of the Max Planck Institute for Plant Breeding Research. For qRT-PCR, complement DNA was synthesized from 5 µg RNA using Super Script® II Reverse Transcriptase in a 20 µl system (Invitrogen). Quantitative RT-PCR was performed on a CFX384 Touch™ Real-Time PCR Detection System using SYBR Green dye (iQ™ SYBR® Green Supermix, Biorad). Expression values were normalized to the expression of *PP2A* and subject to further statistical analysis.

**Table 3-1. Primers used in for qRT-PCR.**

Gene	qPCR-primer-F	qPCR-primer-R
<i>VIH2</i>	ATGGAGATGGAAGAAGGAGC	GCAATCAAGCAATCACAAATGGG
<i>FLC</i>	CCGAACATCATGTTGAAGCTTGTGAG	CGGAGATTTGTCCAGCAGGTG
<i>FT</i>	GGTGGAGAAGACCTCAGGAA	ACCCTGGTGCATACACTGTT
<i>SOC1</i>	ATGAATTCGCCAGCTCCAATAT	GCTTCATATTTCAAATGCTGCA
<i>PP2A</i>	TAACGTGGCCAAAATGATGC	GTTCTCCACAACCGCTTGGT

### 3.4.4 HPLC Analyses of Inositol Phosphates

HPLC analyses were performed by our collaborator, Dr Gabriel Schaaf's group; according to the procedures in their published work (Laha et al., 2015).

### 3.4.5 Root Length Measurement

Seeds were sown on sucrose-free MS agar medium with 0, 10 or 20 MeJA. Seeds were stratified for 5 days and then transferred to an environmental chamber set up to 12-hour photoperiods. Responses to MeJA for each *VIH2*-NIL were obtained by measuring root length in 10-day-old seedlings as well as the number of seedlings with complete root growth inhibition. For seedlings without root, root length was recorded as 0 mm.

### 3.4.6 Analysis on *VIH2* allele variation and distribution

*VIH2* genetic sequence information was downloaded from the 1001 Genomes Tools website as a VCF file. The geographic location of accessions and information on their genetic group

membership are available at: <http://1001genomes.org/tables/1001genomes-accessions.html>. The *VIH2* allele distribution maps is accessible at:

[https://www.google.com/fusiontables/DataSource?docid=1DPa0IxZNgPz18gwdIRgYGWn\\_QIplyj5571SOjtsZ&invite=CK217JsD](https://www.google.com/fusiontables/DataSource?docid=1DPa0IxZNgPz18gwdIRgYGWn_QIplyj5571SOjtsZ&invite=CK217JsD)

### 3.4.7 Accession Numbers

Sequences of VIH homolog proteins were downloaded from online databases. *Chlamydomonas* sequences are available from the Phytozome website under gene IDs *CrVIP1* and, Cre03.g185500 (<https://phytozome.jgi.doe.gov/pz/portal.html>). Other sequences were obtained from Genbank, and their IDs are listed in Table 3-2. Sequence alignment was performed by Clustal Omega on the EMBL-EBI website.

**Table 3-2. The species and GenBank ID of VIP homologs protein sequences used in the alignment.**

<i>Arabidopsis thaliana</i>	Q84WW3.1	<i>Mus musculus</i>	A2ARP1.1
<i>Arabidopsis thaliana</i>	F4J8C6.1	<i>Mus musculus</i>	Q6ZQB6.3
<i>Homo sapiens</i>	Q6PFW1.1	<i>S. cerevisiae</i>	Q06685.1
<i>Homo sapiens</i>	O43314.3	<i>S. pombe</i>	O74429.1
<i>Rattus norvegicus</i>	POC644.1	<i>Aspergillus nidulans</i>	XP_663401.1
		<i>Ustilago maydis</i>	XP_011392692.1
<i>VIH2</i> homolog predicted protein sequence in <i>Brassicaceae</i> species			
<i>Arabidopsis lyrata</i>	XP_002882186.1	<i>Eutrema salsugineum</i>	XP_006408486.1
<i>Capsella rubella</i>	XP_006296593.1	<i>Camelina sativa</i>	XP_010508557.1

**Table 3-3. Primers used for fine mapping of QTL3.**

Marker type	Pos (bp)	Primer-F	Primer-R
Indel	786384	CATCCGAATGCCATTGTTC	AGCTGCTTCCTTATAGCGTCC
Indel	20573	AGAAAACGATTGTTATCGGTGAAA	AACAATCTCCAAAGACACACACAC
Indel	21228	TTTATCCGAACGATCATTTAACAC	GCAAATACTAAAATGAACGTGGAA
Indel	30272	GAAGAGGAAAGAAAATGACGTAGC	CCAGAGAACAACAACACATCAGAG
Indel	40921	AAATCCAAAAGCATCAACAACATA	AGAAAAGGAGTGTGAGTTTCGAGT
Indel	46489	GCAATTCTGCAAAAACGAATACTA	TCTAAGGGTCTTCAACCAAACATT
Indel	50026	CACGTAACGTAGCTGAGATTGAGT	TGACTGACTTGTGTTCTATTCCG
Indel	66389	GAAAACAAACAAAGAAACAAAACG	GCCTCCTCTTTCACCTATTTACTC
Indel	67187	ACAATACCCAAGAAAACAATCACA	GTGTCTCTATGGTGAAAATGGGTC
Indel	75353	ACATGTTTAATGGCATTGATGAAA	GAATTGTGTCTTGATCTTCTCCTG
Indel	75353	ACATGTTTAATGGCATTGATGAAA	GAATTGTGTCTTGATCTTCTCCTG
Indel	78287	AAAAGACACATTGTCATAACCAAA	TCTGATTACAGTTTGGGATACAT
Indel	82141	AGGCGGGTTTAAATCTTCTTAATC	GGTGATTCATGCAGCACTACAATA
Indel	87800	GCCGGATTTGTAGAAGAATAAAAA	GACACAGGTCCACTTACTCAACAC
Indel	93679	ATACGAAGAAGAGAGAAGAGACGG	GTAATGCAATTAGGGAACATTTGTC
CAPs Mbol	97041	CATGGAAGAGTCCTTGCTAACC	TTACAGCTCTGACGAGGGACG
Indel	99101	CCTCAGATCGTAGAAGATCAAACC	GAAAAGCAAATGGCTAGAGAAGTT
indel	101610	CCTTCTTTTCCATGACGCAGA	GGTTTTGATTTTTTGAATCTGTGAC
Indel	102346	TGTGTGTCGTCCATCAATGTTAC	AAAGCAACCAAAAAGCTAAATGAC
CAPs MseI	103890	GTGAAAAAGTGGAAGAACTGTGG	AAACACTAGATGGACGGTTGAGA
Indel	104543	CAAAAGTTTGTGTAACCGTGGG	GTTCAATTAATGGAAGAGGTTT
CAPs BanI	109606	ATGAAGAGAGCTTGTTTCAGATGC	CCAGACCTTATCAAAATCGTCAC
dCAPs SphI	111520	GAGAATGGTGAGCATGATGAGATgcATG	GATAAGAGAACACGTTCTGACTGTGG
Indel	114823	CCTCCGTTTGTGTAAGTCTTATC	AAGAATAATGTTGATTACCCGTGC
Indel	120128	GAAATTAATGTTGTGTAGGGAATGTTT	CACATATTCAATGATTTTCCTATGC
Indel	120128	GAAATTAATGTTGTGTAGGGAATGTTT	CACATATTCAATGATTTTCCTATGC
Indel	122886	CAAGATTACGTTTGAGCCAGTAT	TCAGATCTTGGTACCTTTTGTAC
Indel	124224	ATCCAAGGATAAATGAATGGAAGA	TGTTTTGACTTTTGTACCTTGTGC
Indel	125336	CATTTGCAATGTCGTAAAGTCTTC	TGATCTCATCACACTTGATAGTGC
Indel	125336	CATTTGCAATGTCGTAAAGTCTTC	TGATCTCATCACACTTGATAGTGC
Indel	162163	CTCCGTAACCTCAAAAAGGAAAAA	CGTCGATCAGAAATCAAAGGTAAT
Indel	165037	AGTGCTTGATGATGTCAGTGATTT	TTGCGACTCTTTATTTCTTTGTCA
Indel	184868	TCCGTTGGATTGTAATTATACCTT	TAGATGAATGAGGATCAAGCGTAG
Indel	186431	TGATTGATATGTAACCAATTCCA	TTTCGAGATCATAGGATTTAAGGA
Indel	186638	GAAAAATCTCAAATCGAAGGAAAA	ACTATCCAATGAAAAGATGCCAC
Indel	187046	TCGTCGTCTATGGATGATTTAATG	TTCTTTTCAGAGAACCACACACAT
Indel	190365	CTCAGGTTGTGATTCTTTTCAAC	ATCCTGAGTCATTCTCAAACCTCCT
Indel	190365	CTCAGGTTGTGATTCTTTTCAAC	ATCCTGAGTCATTCTCAAACCTCCT
Indel	193410	CAAAACCAAAACCTTTTCAAGTTA	TTCCCTATTTTCTGCTTCTTTTG

Indel	193504	CAAAAAGAAGCAGAAAATAGGGAA	GTCACGTGTAAACGTTTGTTGTC
Indel	197733	TATACCAATAAGGGGCAATGATA	CTCTCTCTCGTACTTCTCCACC
Indel	201851	AGAAATTAGTGGTCCACGTGTCTT	TTCCATGTAACTTTGTTTCGTGT
Indel	205505	GTGTATGCTCATGAGACCAGAGAG	GAGAAGAAGAAGCTCGTAACAACC
Indel	217885	TGATTGATTCGTCACATTAGAAGC	GCTGGAGTCTCTTAGCATCTTTC
Indel	303845	ATATGATATTTTCGATTCACGGG	ACGGCGTAAATTCTGTGTTGTAA
Indel	333822	CATGAATTAATTTGGATGGGTTTA	GCATTCGTGTATAAAGGACACAA
Indel	354339	ATTCTTCTGGTTGGGAATGTAGTC	ATCATCTTTAACAAGAACCCAGA

**Table 3-4. Primers used in Sanger sequencing of *VIH2* locus. Position of the primer is relative to the start codon of *VIH2*.**

Position	Forward primer	Position	Reverse primer
-2533	GAAGATCAAGAATCCGTTTCATC	-1440	CAATTGAAAAAATTAGGAATGC
-1546	GGAAAGACGATGATGTGCAT	-368	TCTAAAAGCTGCGGTGAAGC
-526	GACCATGATAGTGTATTTAACG	693	AACGAATCTTTCATAACATGAG
613	CCACAGATCTTACATTTTGGG	1698	CTCCTCGTATATATAAGAGCC
1599	ATTGTCATAATTCTGATTAGCTC	2774	CTTACAGTGCAAGAGATTTGTG
2692	CATAGTTGTATTTATAGGTTAAG	3832	ATATCAAAGGTCTAATATGAGAG
3694	TCCTCAATTTATATCACTGCC	4933	ATTCAATCTAGCCTGCAGGG
4817	TGGTCAACAAGATCGAAACTG	6037	TACTCTTATCTTAGATTGTCACC
5919	AGCTAAGTTATGTGTTATGTGG	7190	CACTTTTTGAGAAGTAACCCG
7101	GTTTAAGCTTGGAGACAATGG	8121	GAAGTAAGATTCACCTATTAGGGG

## REFERENCES

- Abulencia, A., Adelman, J., Affolder, T., Akimoto, T., Albrow, M.G., Ambrose, D., Amerio, S., Amidei, D., Anastassov, A., Anikeev, K., *et al.* (2006). Observation of Bs(0)-Bs(0) oscillations. *Phys Rev Lett* **97**, 242003.
- Achard, P., Herr, A., Baulcombe, D.C., and Harberd, N.P. (2004). Modulation of floral development by a gibberellin-regulated microRNA. *Development* **131**, 3357-3365.
- Agrena, J., Oakley, C.G., McKay, J.K., Lovell, J.T., and Schemske, D.W. (2013). Genetic mapping of adaptation reveals fitness tradeoffs in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* **110**, 21077-21082.
- Airoidi, C.A., McKay, M., and Davies, B. (2015). MAF2 Is Regulated by Temperature-Dependent Splicing and Represses Flowering at Low Temperatures in Parallel with FLM. *PLoS One* **10**, e0126516.
- Alonso-Blanco, C., Aarts, M.G., Bentsink, L., Keurentjes, J.J., Reymond, M., Vreugdenhil, D., and Koornneef, M. (2009). What has natural variation taught us about plant development, physiology, and adaptation? *Plant Cell* **21**, 1877-1896.
- Alonso-Blanco, C., Gomez-Mena, C., Llorente, F., Koornneef, M., Salinas, J., and Martinez-Zapater, J.M. (2005). Genetic and molecular analyses of natural variation indicate CBF2 as a candidate gene for underlying a freezing tolerance quantitative trait locus in *Arabidopsis*. *Plant Physiol* **139**, 1304-1312.
- Amasino, R. (2010). Seasonal and developmental timing of flowering. *Plant J* **61**, 1001-1013.
- Angel, A., Song, J., Dean, C., and Howard, M. (2011). A Polycomb-based switch underlying quantitative epigenetic memory. *Nature* **476**, 105-108.
- Atwell, S., Huang, Y.S., Vilhjalmsen, B.J., Willems, G., Horton, M., Li, Y., Meng, D., Platt, A., Tarone, A.M., Hu, T.T., *et al.* (2010). Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* **465**, 627-631.
- Aukerman, M.J., and Sakai, H. (2003). Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. *Plant Cell* **15**, 2730-2741.
- Azpiroz, R., Wu, Y., LoCascio, J.C., and Feldmann, K.A. (1998). An *Arabidopsis* brassinosteroid-dependent mutant is blocked in cell elongation. *Plant Cell* **10**, 219-230.
- Balasubramanian, S., Sureshkumar, S., Lempe, J., and Weigel, D. (2006). Potent induction of *Arabidopsis thaliana* flowering by elevated growth temperature. *PLoS Genet* **2**, e106.
- Bentsink, L., Yuan, K., Koornneef, M., and Vreugdenhil, D. (2003). The genetics of phytate and phosphate accumulation in seeds and leaves of *Arabidopsis thaliana*, using natural variation. *Theor Appl Genet* **106**, 1234-1243.
- Botto, J.F., and Coluccio, M.P. (2007). Seasonal and plant-density dependency for quantitative trait loci affecting flowering time in multiple populations of *Arabidopsis thaliana*. *Plant Cell Environ* **30**, 1465-1479.
- Burghardt, L.T., Runcie, D.E., Wilczek, A.M., Cooper, M.D., Roe, J.L., Welch, S.M., and Schmitt, J. (2016). Fluctuating, warm temperatures decrease the effect of a key floral repressor on flowering time in *Arabidopsis thaliana*. *New Phytol* **210**, 564-576.



- Burn, J.E., Smyth, D.R., Peacock, W.J., and Dennis, E.S. (1993). Genes Conferring Late Flowering in *Arabidopsis-Thaliana*. *Genetica* 90, 147-155.
- Caicedo, A.L., Stinchcombe, J.R., Olsen, K.M., Schmitt, J., and Purugganan, M.D. (2004). Epistatic interaction between *Arabidopsis* FRI and FLC flowering time genes generates a latitudinal cline in a life history trait. *Proc Natl Acad Sci U S A* 101, 15670-15675.
- Chakraborty, A., Koldobskiy, M.A., Bello, N.T., Maxwell, M., Potter, J.J., Juluri, K.R., Maag, D., Kim, S., Huang, A.S., Dailey, M.J., *et al.* (2010). Inositol pyrophosphates inhibit Akt signaling, thereby regulating insulin sensitivity and weight gain. *Cell* 143, 897-910.
- Chandler, J., Martinez-Zapater, J.M., and Dean, C. (2000). Mutations causing defects in the biosynthesis and response to gibberellins, abscisic acid and phytochrome B do not inhibit vernalization in *Arabidopsis fca-1*. *Planta* 210, 677-682.
- Chen, X. (2004). A microRNA as a translational repressor of APETALA2 in *Arabidopsis* flower development. *Science* 303, 2022-2025.
- Chini, A., Fonseca, S., Fernandez, G., Adie, B., Chico, J.M., Lorenzo, O., Garcia-Casado, G., Lopez-Vidriero, I., Lozano, F.M., Ponce, M.R., *et al.* (2007). The JAZ family of repressors is the missing link in jasmonate signalling. *Nature* 448, 666-671.
- Cho, H.J., Kim, J.J., Lee, J.H., Kim, W., Jung, J.H., Park, C.M., and Ahn, J.H. (2012). SHORT VEGETATIVE PHASE (SVP) protein negatively regulates miR172 transcription via direct binding to the pri-miR172a promoter in *Arabidopsis*. *FEBS Lett* 586, 2332-2337.
- Choi, J., Hyun, Y., Kang, M.J., In Yun, H., Yun, J.Y., Lister, C., Dean, C., Amasino, R.M., Noh, B., Noh, Y.S., *et al.* (2009). Resetting and regulation of Flowering Locus C expression during *Arabidopsis* reproductive development. *Plant J* 57, 918-931.
- Choi, K., Kim, J., Hwang, H.J., Kim, S., Park, C., Kim, S.Y., and Lee, I. (2011). The FRIGIDA complex activates transcription of FLC, a strong flowering repressor in *Arabidopsis*, by recruiting chromatin modification factors. *Plant Cell* 23, 289-303.
- Chory, J., Nagpal, P., and Peto, C.A. (1991). Phenotypic and Genetic Analysis of *det2*, a New Mutant That Affects Light-Regulated Seedling Development in *Arabidopsis*. *Plant Cell* 3, 445-459.
- Clarke, J.H., and Dean, C. (1994). Mapping FRI, a locus controlling flowering time and vernalization response in *Arabidopsis thaliana*. *Mol Gen Genet* 242, 81-89.
- Clarke, J.H., Mithen, R., Brown, J.K.M., and Dean, C. (1995). Qtl Analysis of Flowering Time in *Arabidopsis-Thaliana*. *Molecular & General Genetics* 248, 278-286.
- Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J* 16, 735-743.
- Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q., Searle, I., Giakountis, A., Farrona, S., Gissot, L., Turnbull, C., *et al.* (2007). FT protein movement contributes to long-distance signaling in floral induction of *Arabidopsis*. *Science* 316, 1030-1033.
- Couso, I., Evans, B., Li, J., Liu, Y., Ma, F., Diamond, S., Allen, D.K., and Umen, J.G. (2016). Synergism between inositol polyphosphates and TOR kinase signaling in nutrient sensing, growth control and lipid metabolism in *Chlamydomonas*. *Plant Cell*.
- Davies, P.J. (2004). *Plant hormones : biosynthesis, signal transduction, action!*, 3rd edn (Dordrecht ; Boston: Kluwer Academic).
- Davis, S.J. (2009). Integrating hormones into the floral-transition pathway of *Arabidopsis thaliana*. *Plant Cell Environ* 32, 1201-1210.

- De Lucia, F., Crevillen, P., Jones, A.M., Greb, T., and Dean, C. (2008). A PHD-polycomb repressive complex 2 triggers the epigenetic silencing of FLC during vernalization. *Proc Natl Acad Sci U S A* *105*, 16831-16836.
- Desai, M., Rangarajan, P., Donahue, J.L., Williams, S.P., Land, E.S., Mandal, M.K., Phillippy, B.Q., Perera, I.Y., Raboy, V., and Gillaspay, G.E. (2014). Two inositol hexakisphosphate kinases drive inositol pyrophosphate synthesis in plants. *Plant J* *80*, 642-653.
- Ding, L., Kim, S.Y., and Michaels, S.D. (2013). FLOWERING LOCUS C EXPRESSOR family proteins regulate FLOWERING LOCUS C expression in both winter-annual and rapid-cycling Arabidopsis. *Plant Physiol* *163*, 243-252.
- Domagalska, M.A., Schomburg, F.M., Amasino, R.M., Vierstra, R.D., Nagy, F., and Davis, S.J. (2007). Attenuation of brassinosteroid signaling enhances FLC expression and delays flowering. *Development* *134*, 2841-2850.
- Doyle, M.R., Bizzell, C.M., Keller, M.R., Michaels, S.D., Song, J., Noh, Y.S., and Amasino, R.M. (2005). HUA2 is required for the expression of floral repressors in Arabidopsis thaliana. *Plant J* *41*, 376-385.
- el-Lithy, M.E., Bentsink, L., Hanhart, C.J., Ruys, G.J., Rovito, D., Broekhof, J.L., van der Poel, H.J., van Eijk, M.J., Vreugdenhil, D., and Koornneef, M. (2006). New Arabidopsis recombinant inbred line populations genotyped using SNPWave and their use for mapping flowering-time quantitative trait loci. *Genetics* *172*, 1867-1876.
- El-Lithy, M.E., Clerckx, E.J., Ruys, G.J., Koornneef, M., and Vreugdenhil, D. (2004). Quantitative trait locus analysis of growth-related traits in a new Arabidopsis recombinant inbred population. *Plant Physiol* *135*, 444-458.
- Fernandez, V., Takahashi, Y., Le Gourrierc, J., and Coupland, G. (2016). Photoperiodic and thermosensory pathways interact through CONSTANS to promote flowering at high temperature under short days. *Plant J* *86*, 426-440.
- Fornara, F., Panigrahi, K.C., Gissot, L., Sauerbrunn, N., Ruhl, M., Jarillo, J.A., and Coupland, G. (2009). Arabidopsis DOF transcription factors act redundantly to reduce CONSTANS expression and are essential for a photoperiodic flowering response. *Dev Cell* *17*, 75-86.
- Fournier-Level, A., Korte, A., Cooper, M.D., Nordborg, M., Schmitt, J., and Wilczek, A.M. (2011). A Map of Local Adaptation in Arabidopsis thaliana. *Science* *334*, 86-89.
- Galvao, V.C., Collani, S., Horrer, D., and Schmid, M. (2015). Gibberellic acid signaling is required for ambient temperature-mediated induction of flowering in Arabidopsis thaliana. *Plant J* *84*, 949-962.
- Galvao, V.C., Horrer, D., Kuttner, F., and Schmid, M. (2012). Spatial control of flowering by DELLA proteins in Arabidopsis thaliana. *Development* *139*, 4072-4082.
- Gan, X., Stegle, O., Behr, J., Steffen, J.G., Drewe, P., Hildebrand, K.L., Lyngsoe, R., Schultheiss, S.J., Osborne, E.J., Sreedharan, V.T., *et al.* (2011). Multiple reference genomes and transcriptomes for Arabidopsis thaliana. *Nature* *477*, 419-423.
- Gandikota, M., Birkenbihl, R.P., Hohmann, S., Cardon, G.H., Saedler, H., and Huijser, P. (2007). The miRNA156/157 recognition element in the 3' UTR of the Arabidopsis SBP box gene SPL3 prevents early flowering by translational inhibition in seedlings. *Plant J* *49*, 683-693.
- Gazzani, S., Gendall, A.R., Lister, C., and Dean, C. (2003). Analysis of the molecular basis of flowering time variation in Arabidopsis accessions. *Plant Physiol* *132*, 1107-1114.

- Geraldo, N., Baurle, I., Kidou, S., Hu, X., and Dean, C. (2009). FRIGIDA delays flowering in *Arabidopsis* via a cotranscriptional mechanism involving direct interaction with the nuclear cap-binding complex. *Plant Physiol* 150, 1611-1618.
- Ghandilyan, A., Ilk, N., Hanhart, C., Mbengue, M., Barboza, L., Schat, H., Koornneef, M., El-Lithy, M., Vreugdenhil, D., Reymond, M., *et al.* (2009). A strong effect of growth medium and organ type on the identification of QTLs for phytate and mineral concentrations in three *Arabidopsis thaliana* RIL populations. *J Exp Bot* 60, 1409-1425.
- Gilliland, L.U., Magallanes-Lundback, M., Hemming, C., Supplee, A., Koornneef, M., Bentsink, L., and Dellapenna, D. (2006). Genetic basis for natural variation in seed vitamin E levels in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* 103, 18834-18841.
- Granier, C., Aguirrezabal, L., Chenu, K., Cookson, S.J., Dauzat, M., Hamard, P., Thioux, J.J., Rolland, G., Bouchier-Combaud, S., Lebaudy, A., *et al.* (2006). PHENOPSIS, an automated platform for reproducible phenotyping of plant responses to soil water deficit in *Arabidopsis thaliana* permitted the identification of an accession with low sensitivity to soil water deficit. *New Phytol* 169, 623-635.
- Griffiths, J., Murase, K., Rieu, I., Zentella, R., Zhang, Z.L., Powers, S.J., Gong, F., Phillips, A.L., Hedden, P., Sun, T.P., *et al.* (2006). Genetic characterization and functional analysis of the GID1 gibberellin receptors in *Arabidopsis*. *Plant Cell* 18, 3399-3414.
- Grillo, M.A., Li, C., Hammond, M., Wang, L., and Schemske, D.W. (2013). Genetic architecture of flowering time differentiation between locally adapted populations of *Arabidopsis thaliana*. *New Phytol* 197, 1321-1331.
- Gross, D., Wilczek, F., Hooft, G., Deser, S., Belinski, V., and Nilsson, B. (2009). Recorded in Santa Barbara. *Quantum Mechanics of Fundamental Systems: The Quest for Beauty and Simplicity*, Claudio Bunster Festschrift, 3-+.
- Hagenblad, J., Tang, C., Molitor, J., Werner, J., Zhao, K., Zheng, H., Marjoram, P., Weigel, D., and Nordborg, M. (2004). Haplotype structure and phenotypic associations in the chromosomal regions surrounding two *Arabidopsis thaliana* flowering time loci. *Genetics* 168, 1627-1638.
- Hayama, R., Sarid-Krebs, L., Richter, R., Fernandez, V., Jang, S., and Coupland, G. (2017). PSEUDO RESPONSE REGULATORS stabilize CONSTANS protein to promote flowering in response to day length. *EMBO J* 36, 904-918.
- Heo, J.B., and Sung, S. (2011). Encoding memory of winter by noncoding RNAs. *Epigenetics* 6, 544-547.
- Horniyk, C., Terzi, L.C., and Simpson, G.G. (2010). The spen family protein FPA controls alternative cleavage and polyadenylation of RNA. *Dev Cell* 18, 203-213.
- Hu, X., Kong, X., Wang, C., Ma, L., Zhao, J., Wei, J., Zhang, X., Loake, G.J., Zhang, T., Huang, J., *et al.* (2014). Proteasome-mediated degradation of FRIGIDA modulates flowering time in *Arabidopsis* during vernalization. *Plant Cell* 26, 4763-4781.
- Huang, X., Paulo, M.J., Boer, M., Effgen, S., Keizer, P., Koornneef, M., and van Eeuwijk, F.A. (2011). Analysis of natural allelic variation in *Arabidopsis* using a multiparent recombinant inbred line population. *Proc Natl Acad Sci U S A* 108, 4488-4493.
- Hyun, K.G., Oh, J.E., Park, J., Noh, Y.S., and Song, J.J. (2016). Structural Analysis of FRIGIDA Flowering-Time Regulator. *Mol Plant* 9, 618-620.
- Ilk, N., Ding, J., Ihnatowicz, A., Koornneef, M., and Reymond, M. (2015). Natural variation for anthocyanin accumulation under high-light and low-temperature stress is attributable to the

- ENHANCER OF AG-4 2 (HUA2) locus in combination with PRODUCTION OF ANTHOCYANIN PIGMENT1 (PAP1) and PAP2. *New Phytol* 206, 422-435.
- Irwin, J.A., Soumpourou, E., Lister, C., Lighthart, J.D., Kennedy, S., and Dean, C. (2016). Nucleotide polymorphism affecting FLC expression underpins heading date variation in horticultural brassicas. *Plant J* 87, 597-605.
- Johanson, U., West, J., Lister, C., Michaels, S., Amasino, R., and Dean, C. (2000). Molecular analysis of FRIGIDA, a major determinant of natural variation in Arabidopsis flowering time. *Science* 290, 344-347.
- Jung, J.H., Ju, Y., Seo, P.J., Lee, J.H., and Park, C.M. (2012). The SOC1-SPL module integrates photoperiod and gibberellic acid signals to control flowering time in Arabidopsis. *Plant J* 69, 577-588.
- Katsir, L., Schilmiller, A.L., Staswick, P.E., He, S.Y., and Howe, G.A. (2008). COI1 is a critical component of a receptor for jasmonate and the bacterial virulence factor coronatine. *Proc Natl Acad Sci U S A* 105, 7100-7105.
- Kim, J.J., Lee, J.H., Kim, W., Jung, H.S., Huijser, P., and Ahn, J.H. (2012). The microRNA156-SQUAMOSA PROMOTER BINDING PROTEIN-LIKE3 module regulates ambient temperature-responsive flowering via FLOWERING LOCUS T in Arabidopsis. *Plant Physiol* 159, 461-478.
- Kim, S., Choi, K., Park, C., Hwang, H.J., and Lee, I. (2006). SUPPRESSOR OF FRIGIDA4, encoding a C2H2-Type zinc finger protein, represses flowering by transcriptional activation of Arabidopsis FLOWERING LOCUS C. *Plant Cell* 18, 2985-2998.
- Kim, S.Y., Yu, X., and Michaels, S.D. (2008). Regulation of CONSTANS and FLOWERING LOCUS T expression in response to changing light quality. *Plant Physiol* 148, 269-279.
- Kim, W.Y., Fujiwara, S., Suh, S.S., Kim, J., Kim, Y., Han, L., David, K., Putterill, J., Nam, H.G., and Somers, D.E. (2007). ZEITLUPE is a circadian photoreceptor stabilized by GIGANTEA in blue light. *Nature* 449, 356-360.
- Ko, J.H., Mitina, I., Tamada, Y., Hyun, Y., Choi, Y., Amasino, R.M., Noh, B., and Noh, Y.S. (2010). Growth habit determination by the balance of histone methylation activities in Arabidopsis. *EMBO J* 29, 3208-3215.
- Kobayashi, Y., and Weigel, D. (2007). Move on up, it's time for change--mobile signals controlling photoperiod-dependent flowering. *Genes Dev* 21, 2371-2384.
- Koncz, C., Nemeth, K., Redei, G.P., and Schell, J. (1992). T-DNA insertional mutagenesis in Arabidopsis. *Plant Mol Biol* 20, 963-976.
- Koornneef, M., Alonso-Blanco, C., and Vreugdenhil, D. (2004). Naturally occurring genetic variation in Arabidopsis thaliana. *Annu Rev Plant Biol* 55, 141-172.
- Koornneef, M., Hanhart, C.J., and van der Veen, J.H. (1991). A genetic and physiological analysis of late flowering mutants in Arabidopsis thaliana. *Mol Gen Genet* 229, 57-66.
- Kover, P.X., Valdar, W., Trakalo, J., Scarcelli, N., Ehrenreich, I.M., Purugganan, M.D., Durrant, C., and Mott, R. (2009). A Multiparent Advanced Generation Inter-Cross to fine-map quantitative traits in Arabidopsis thaliana. *PLoS Genet* 5, e1000551.
- Kumar, S.V., Lucyshyn, D., Jaeger, K.E., Alos, E., Alvey, E., Harberd, N.P., and Wigge, P.A. (2012). Transcription factor PIF4 controls the thermosensory activation of flowering. *Nature* 484, 242-245.

- Kuo, H.F., Chang, T.Y., Chiang, S.F., Wang, W.D., Charng, Y.Y., and Chiou, T.J. (2014). Arabidopsis inositol pentakisphosphate 2-kinase, AtIPK1, is required for growth and modulates phosphate homeostasis at the transcriptional level. *Plant J* 80, 503-515.
- Laha, D., Johnen, P., Azevedo, C., Dynowski, M., Weiss, M., Capolicchio, S., Mao, H., Iven, T., Steenbergen, M., Freyer, M., *et al.* (2015). VIH2 Regulates the Synthesis of Inositol Pyrophosphate InsP8 and Jasmonate-Dependent Defenses in Arabidopsis. *Plant Cell* 27, 1082-1097.
- Laubinger, S., Marchal, V., Le Gourrierc, J., Wenkel, S., Adrian, J., Jang, S., Kulajta, C., Braun, H., Coupland, G., and Hoecker, U. (2006). Arabidopsis SPA proteins regulate photoperiodic flowering and interact with the floral inducer CONSTANS to regulate its stability. *Development* 133, 3213-3222.
- Lazaro, A., Valverde, F., Pineiro, M., and Jarillo, J.A. (2012). The Arabidopsis E3 ubiquitin ligase HOS1 negatively regulates CONSTANS abundance in the photoperiodic control of flowering. *Plant Cell* 24, 982-999.
- Le Corre, V., Roux, F., and Reboud, X. (2002). DNA polymorphism at the FRIGIDA gene in Arabidopsis thaliana: extensive nonsynonymous variation is consistent with local selection for flowering time. *Mol Biol Evol* 19, 1261-1271.
- Lee, I., Aukerman, M.J., Gore, S.L., Lohman, K.N., Michaels, S.D., Weaver, L.M., John, M.C., Feldmann, K.A., and Amasino, R.M. (1994). Isolation of LUMINIDEPENDENS: a gene involved in the control of flowering time in Arabidopsis. *Plant Cell* 6, 75-83.
- Lee, I., Bleecker, A., and Amasino, R. (1993). Analysis of naturally occurring late flowering in Arabidopsis thaliana. *Mol Gen Genet* 237, 171-176.
- Lee, J.H., Ryu, H.S., Chung, K.S., Pose, D., Kim, S., Schmid, M., and Ahn, J.H. (2013). Regulation of temperature-responsive flowering by MADS-box transcription factor repressors. *Science* 342, 628-632.
- Lee, Y.S., Mulugu, S., York, J.D., and O'Shea, E.K. (2007). Regulation of a cyclin-CDK-CDK inhibitor complex by inositol pyrophosphates. *Science* 316, 109-112.
- Lempe, J., Balasubramanian, S., Sureshkumar, S., Singh, A., Schmid, M., and Weigel, D. (2005). Diversity of flowering responses in wild Arabidopsis thaliana strains. *PLoS Genet* 1, 109-118.
- Li, D., Liu, C., Shen, L., Wu, Y., Chen, H., Robertson, M., Helliwell, C.A., Ito, T., Meyerowitz, E., and Yu, H. (2008). A repressor complex governs the integration of flowering signals in Arabidopsis. *Dev Cell* 15, 110-120.
- Li, P., Filiault, D., Box, M.S., Kerdaffrec, E., van Oosterhout, C., Wilczek, A.M., Schmitt, J., McMullan, M., Bergelson, J., Nordborg, M., *et al.* (2014). Multiple FLC haplotypes defined by independent cis-regulatory variation underpin life history diversity in Arabidopsis thaliana. *Genes Dev* 28, 1635-1640.
- Lim, M.H., Kim, J., Kim, Y.S., Chung, K.S., Seo, Y.H., Lee, I., Kim, J., Hong, C.B., Kim, H.J., and Park, C.M. (2004). A new Arabidopsis gene, FLK, encodes an RNA binding protein with K homology motifs and regulates flowering time via FLOWERING LOCUS C. *Plant Cell* 16, 731-740.
- Liu, C., Chen, H., Er, H.L., Soo, H.M., Kumar, P.P., Han, J.H., Liou, Y.C., and Yu, H. (2008a). Direct interaction of AGL24 and SOC1 integrates flowering signals in Arabidopsis. *Development* 135, 1481-1491.
- Liu, F., Marquardt, S., Lister, C., Swiezewski, S., and Dean, C. (2010). Targeted 3' processing of antisense transcripts triggers Arabidopsis FLC chromatin silencing. *Science* 327, 94-97.

- Liu, F., Quesada, V., Crevillen, P., Baurle, I., Swiezewski, S., and Dean, C. (2007). The Arabidopsis RNA-binding protein FCA requires a lysine-specific demethylase 1 homolog to downregulate FLC. *Mol Cell* 28, 398-407.
- Liu, L.J., Zhang, Y.C., Li, Q.H., Sang, Y., Mao, J., Lian, H.L., Wang, L., and Yang, H.Q. (2008b). COP1-mediated ubiquitination of CONSTANS is implicated in cryptochrome regulation of flowering in Arabidopsis. *Plant Cell* 20, 292-306.
- Macknight, R., Bancroft, I., Page, T., Lister, C., Schmidt, R., Love, K., Westphal, L., Murphy, G., Sherson, S., Cobbett, C., *et al.* (1997). FCA, a gene controlling flowering time in Arabidopsis, encodes a protein containing RNA-binding domains. *Cell* 89, 737-745.
- McClung, C.R., Lou, P., Hermand, V., and Kim, J.A. (2016). The Importance of Ambient Temperature to Growth and the Induction of Flowering. *Frontiers in plant science* 7, 1266.
- Mendez-Vigo, B., Pico, F.X., Ramiro, M., Martinez-Zapater, J.M., and Alonso-Blanco, C. (2011). Altitudinal and climatic adaptation is mediated by flowering traits and FRI, FLC, and PHYC genes in Arabidopsis. *Plant Physiol* 157, 1942-1955.
- Menniti, F.S., Miller, R.N., Putney, J.W., Jr., and Shears, S.B. (1993). Turnover of inositol polyphosphate pyrophosphates in pancreatoma cells. *J Biol Chem* 268, 3850-3856.
- Michaels, S.D., and Amasino, R.M. (1999). FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell* 11, 949-956.
- Michaels, S.D., and Amasino, R.M. (2001). Loss of FLOWERING LOCUS C activity eliminates the late-flowering phenotype of FRIGIDA and autonomous pathway mutations but not responsiveness to vernalization. *Plant Cell* 13, 935-941.
- Michaels, S.D., He, Y., Scortecci, K.C., and Amasino, R.M. (2003). Attenuation of FLOWERING LOCUS C activity as a mechanism for the evolution of summer-annual flowering behavior in Arabidopsis. *Proc Natl Acad Sci U S A* 100, 10102-10107.
- Mulugu, S., Bai, W., Fridy, P.C., Bastidas, R.J., Otto, J.C., Dollins, D.E., Haystead, T.A., Ribeiro, A.A., and York, J.D. (2007). A conserved family of enzymes that phosphorylate inositol hexakisphosphate. *Science* 316, 106-109.
- Murase, K., Hirano, Y., Sun, T.P., and Hakoshima, T. (2008). Gibberellin-induced DELLA recognition by the gibberellin receptor GID1. *Nature* 456, 459-463.
- Nakamichi, N., Kita, M., Niinuma, K., Ito, S., Yamashino, T., Mizoguchi, T., and Mizuno, T. (2007). Arabidopsis clock-associated pseudo-response regulators PRR9, PRR7 and PRR5 coordinately and positively regulate flowering time through the canonical CONSTANS-dependent photoperiodic pathway. *Plant & cell physiology* 48, 822-832.
- Napoli, C., Lemieux, C., and Jorgensen, R. (1990). Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell* 2, 279-289.
- Nishizawa, M., Komai, T., Morohashi, N., Shimizu, M., and Toh-e, A. (2008). Transcriptional repression by the Pho4 transcription factor controls the timing of SNZ1 expression. *Eukaryot Cell* 7, 949-957.
- Norbis, F., Boll, M., Stange, G., Markovich, D., Verrey, F., Biber, J., and Murer, H. (1997). Identification of a cDNA/protein leading to an increased Pi-uptake in Xenopus laevis oocytes. *J Membr Biol* 156, 19-24.
- Pigliucci, M., Pollard, H., and Cruzan, M.B. (2003). Comparative studies of evolutionary responses to light environments in Arabidopsis. *Am Nat* 161, 68-82.

- Pohlmann, J., Risse, C., Seidel, C., Pohlmann, T., Jakopiec, V., Walla, E., Ramrath, P., Takeshita, N., Baumann, S., Feldbrugge, M., *et al.* (2014). The Vip1 inositol polyphosphate kinase family regulates polarized growth and modulates the microtubule cytoskeleton in fungi. *PLoS Genet* 10, e1004586.
- Pose, D., Verhage, L., Ott, F., Yant, L., Mathieu, J., Angenent, G.C., Immink, R.G., and Schmid, M. (2013). Temperature-dependent regulation of flowering by antagonistic FLM variants. *Nature* 503, 414-417.
- Prasad, A., Jia, Y., Chakraborty, A., Li, Y., Jain, S.K., Zhong, J., Roy, S.G., Loison, F., Mondal, S., Sakai, J., *et al.* (2011). Inositol hexakisphosphate kinase 1 regulates neutrophil function in innate immunity by inhibiting phosphatidylinositol-(3,4,5)-trisphosphate signaling. *Nat Immunol* 12, 752-760.
- Pulloor, N.K., Nair, S., McCaffrey, K., Kostic, A.D., Bist, P., Weaver, J.D., Riley, A.M., Tyagi, R., Uchil, P.D., York, J.D., *et al.* (2014). Human genome-wide RNAi screen identifies an essential role for inositol pyrophosphates in Type-I interferon response. *PLoS pathogens* 10, e1003981.
- Raboy, V., and Dickinson, D.B. (1987). The Timing and Rate of Phytic Acid Accumulation in Developing Soybean Seeds. *Plant Physiology* 85, 841-844.
- Raboy, V., Dickinson, D.B., and Below, F.E. (1984). Variation in Seed Total Phosphorus, Phytic Acid, Zinc, Calcium, Magnesium, and Protein among Lines of Glycine-Max and Glycine-Soja. *Crop Sci* 24, 431-434.
- Risk, J.M., Laurie, R.E., Macknight, R.C., and Day, C.L. (2010). FRIGIDA and related proteins have a conserved central domain and family specific N- and C- terminal regions that are functionally important. *Plant Mol Biol* 73, 493-505.
- Roldan, M., Gomez-Mena, C., Ruiz-Garcia, L., Salinas, J., and Martinez-Zapater, J.M. (1999). Sucrose availability on the aerial part of the plant promotes morphogenesis and flowering of Arabidopsis in the dark. *Plant J* 20, 581-590.
- Saiardi, A., Erdjument-Bromage, H., Snowman, A.M., Tempst, P., and Snyder, S.H. (1999). Synthesis of diphosphoinositol pentakisphosphate by a newly identified family of higher inositol polyphosphate kinases. *Curr Biol* 9, 1323-1326.
- Saiardi, A., Nagata, E., Luo, H.R., Snowman, A.M., and Snyder, S.H. (2001). Identification and characterization of a novel inositol hexakisphosphate kinase. *J Biol Chem* 276, 39179-39185.
- Salvi, S., and Tuberosa, R. (2005). To clone or not to clone plant QTLs: present and future challenges. *Trends in plant science* 10, 297-304.
- Sawa, M., Nusinow, D.A., Kay, S.A., and Imaizumi, T. (2007). FKF1 and GIGANTEA complex formation is required for day-length measurement in Arabidopsis. *Science* 318, 261-265.
- Schaffer, R., Ramsay, N., Samach, A., Corden, S., Putterill, J., Carre, I.A., and Coupland, G. (1998). The late elongated hypocotyl mutation of Arabidopsis disrupts circadian rhythms and the photoperiodic control of flowering. *Cell* 93, 1219-1229.
- Schlappi, M.R. (2006). FRIGIDA LIKE 2 is a functional allele in Landsberg erecta and compensates for a nonsense allele of FRIGIDA LIKE 1. *Plant Physiol* 142, 1728-1738.
- Schmid, M., Uhlenhaut, N.H., Godard, F., Demar, M., Bressan, R., Weigel, D., and Lohmann, J.U. (2003). Dissection of floral induction pathways using global expression analysis. *Development* 130, 6001-6012.
- Schwab, R., Palatnik, J.F., Riester, M., Schommer, C., Schmid, M., and Weigel, D. (2005). Specific effects of microRNAs on the plant transcriptome. *Dev Cell* 8, 517-527.

- Sheard, L.B., Tan, X., Mao, H., Withers, J., Ben-Nissan, G., Hinds, T.R., Kobayashi, Y., Hsu, F.F., Sharon, M., Browse, J., *et al.* (2010). Jasmonate perception by inositol-phosphate-potentiated COI1-JAZ co-receptor. *Nature* 468, 400-405.
- Sheldon, C.C., Hills, M.J., Lister, C., Dean, C., Dennis, E.S., and Peacock, W.J. (2008). Resetting of FLOWERING LOCUS C expression after epigenetic repression by vernalization. *Proc Natl Acad Sci U S A* 105, 2214-2219.
- Shindo, C., Aranzana, M.J., Lister, C., Baxter, C., Nicholls, C., Nordborg, M., and Dean, C. (2005). Role of FRIGIDA and FLOWERING LOCUS C in determining variation in flowering time of Arabidopsis. *Plant Physiol* 138, 1163-1173.
- Shu, K., Chen, Q., Wu, Y., Liu, R., Zhang, H., Wang, S., Tang, S., Yang, W., and Xie, Q. (2016). ABSCISIC ACID-INSENSITIVE 4 negatively regulates flowering through directly promoting Arabidopsis FLOWERING LOCUS C transcription. *J Exp Bot* 67, 195-205.
- Simon, M., Loudet, O., Durand, S., Berard, A., Brunel, D., Sennesal, F.X., Durand-Tardif, M., Pelletier, G., and Camilleri, C. (2008). Quantitative trait loci mapping in five new large recombinant inbred line populations of Arabidopsis thaliana genotyped with consensus single-nucleotide polymorphism markers. *Genetics* 178, 2253-2264.
- Simpson, G.G. (2004). The autonomous pathway: epigenetic and post-transcriptional gene regulation in the control of Arabidopsis flowering time. *Curr Opin Plant Biol* 7, 570-574.
- Simpson, G.G., Dijkwel, P.P., Quesada, V., Henderson, I., and Dean, C. (2003). FY is an RNA 3' end-processing factor that interacts with FCA to control the Arabidopsis floral transition. *Cell* 113, 777-787.
- Song, J., Angel, A., Howard, M., and Dean, C. (2012a). Vernalization - a cold-induced epigenetic switch. *J Cell Sci* 125, 3723-3731.
- Song, S., Qi, T., Fan, M., Zhang, X., Gao, H., Huang, H., Wu, D., Guo, H., and Xie, D. (2013). The bHLH subgroup IIIId factors negatively regulate jasmonate-mediated plant defense and development. *PLoS Genet* 9, e1003653.
- Song, Y.H., Estrada, D.A., Johnson, R.S., Kim, S.K., Lee, S.Y., MacCoss, M.J., and Imaizumi, T. (2014). Distinct roles of FKF1, Gigantea, and Zeitlupe proteins in the regulation of Constans stability in Arabidopsis photoperiodic flowering. *Proc Natl Acad Sci U S A* 111, 17672-17677.
- Song, Y.H., Smith, R.W., To, B.J., Millar, A.J., and Imaizumi, T. (2012b). FKF1 conveys timing information for CONSTANS stabilization in photoperiodic flowering. *Science* 336, 1045-1049.
- Staswick, P.E., Su, W., and Howell, S.H. (1992). Methyl jasmonate inhibition of root growth and induction of a leaf protein are decreased in an Arabidopsis thaliana mutant. *Proc Natl Acad Sci U S A* 89, 6837-6840.
- Stephens, L., Radenberg, T., Thiel, U., Vogel, G., Khoo, K.H., Dell, A., Jackson, T.R., Hawkins, P.T., and Mayr, G.W. (1993). The detection, purification, structural characterization, and metabolism of diphosphoinositol pentakisphosphate(s) and bisdiphosphoinositol tetrakisphosphate(s). *J Biol Chem* 268, 4009-4015.
- Stinchcombe, J.R., Weinig, C., Ungerer, M., Olsen, K.M., Mays, C., Halldorsdottir, S.S., Purugganan, M.D., and Schmitt, J. (2004). A latitudinal cline in flowering time in Arabidopsis thaliana modulated by the flowering time gene FRIGIDA. *Proc Natl Acad Sci U S A* 101, 4712-4717.
- Streb, H., Irvine, R.F., Berridge, M.J., and Schulz, I. (1983). Release of Ca<sup>2+</sup> from a nonmitochondrial intracellular store in pancreatic acinar cells by inositol-1,4,5-trisphosphate. *Nature* 306, 67-69.



- Suarez-Lopez, P., Wheatley, K., Robson, F., Onouchi, H., Valverde, F., and Coupland, G. (2001). CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. *Nature* 410, 1116-1120.
- Swiezewski, S., Liu, F., Magusin, A., and Dean, C. (2009). Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target. *Nature* 462, 799-802.
- Szjgyarto, Z., Garedew, A., Azevedo, C., and Saiardi, A. (2011). Influence of inositol pyrophosphates on cellular energy dynamics. *Science* 334, 802-805.
- Talbert, P.B., and Henikoff, S. (2014). Environmental responses mediated by histone variants. *Trends Cell Biol* 24, 642-650.
- Tan, X., Calderon-Villalobos, L.I., Sharon, M., Zheng, C., Robinson, C.V., Estelle, M., and Zheng, N. (2007). Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature* 446, 640-645.
- Teotia, S., and Tang, G. (2015). To bloom or not to bloom: role of microRNAs in plant flowering. *Mol Plant* 8, 359-377.
- Thota, S.G., and Bhandari, R. (2015). The emerging roles of inositol pyrophosphates in eukaryotic cell physiology. *J Biosci* 40, 593-605.
- Tisne, S., Reymond, M., Vile, D., Fabre, J., Dauzat, M., Koornneef, M., and Granier, C. (2008). Combined genetic and modeling approaches reveal that epidermal cell area and number in leaves are controlled by leaf and plant developmental processes in Arabidopsis. *Plant Physiology* 148, 1117-1127.
- Tisne, S., Schmalenbach, I., Reymond, M., Dauzat, M., Pervent, M., Vile, D., and Granier, C. (2010). Keep on growing under drought: genetic and developmental bases of the response of rosette area using a recombinant inbred line population. *Plant Cell Environ* 33, 1875-1887.
- Toomajian, C., Hu, T.T., Aranzana, M.J., Lister, C., Tang, C., Zheng, H., Zhao, K., Calabrese, P., Dean, C., and Nordborg, M. (2006). A nonparametric test reveals selection for rapid flowering in the Arabidopsis genome. *PLoS Biol* 4, e137.
- Torti, S., Fornara, F., Vincent, C., Andres, F., Nordstrom, K., Gobel, U., Knoll, D., Schoof, H., and Coupland, G. (2012). Analysis of the Arabidopsis shoot meristem transcriptome during floral transition identifies distinct regulatory patterns and a leucine-rich repeat protein that promotes flowering. *Plant Cell* 24, 444-462.
- Turk, E.M., Fujioka, S., Seto, H., Shimada, Y., Takatsuto, S., Yoshida, S., Wang, H., Torres, Q.I., Ward, J.M., Murthy, G., *et al.* (2005). BAS1 and SOB7 act redundantly to modulate Arabidopsis photomorphogenesis via unique brassinosteroid inactivation mechanisms. *Plant J* 42, 23-34.
- Valverde, F., Mouradov, A., Soppe, W., Ravenscroft, D., Samach, A., and Coupland, G. (2004). Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. *Science* 303, 1003-1006.
- Wahl, V., Ponnu, J., Schlereth, A., Arrivault, S., Langenecker, T., Franke, A., Feil, R., Lunn, J.E., Stitt, M., and Schmid, M. (2013). Regulation of flowering by trehalose-6-phosphate signaling in Arabidopsis thaliana. *Science* 339, 704-707.
- Wang, J.W., Czech, B., and Weigel, D. (2009). miR156-regulated SPL transcription factors define an endogenous flowering pathway in Arabidopsis thaliana. *Cell* 138, 738-749.
- Wang, Z.Y., and Tobin, E.M. (1998). Constitutive expression of the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) gene disrupts circadian rhythms and suppresses its own expression. *Cell* 93, 1207-1217.

- Wei, S., Gruber, M.Y., Yu, B., Gao, M.J., Khachatourians, G.G., Hegedus, D.D., Parkin, I.A., and Hannoufa, A. (2012). Arabidopsis mutant sk156 reveals complex regulation of SPL15 in a miR156-controlled gene network. *BMC Plant Biol* 12, 169.
- Werner, J.D., Borevitz, J.O., Uhlenhaut, N.H., Ecker, J.R., Chory, J., and Weigel, D. (2005). FRIGIDA-independent variation in flowering time of natural Arabidopsis thaliana accessions. *Genetics* 170, 1197-1207.
- Wigge, P.A., Kim, M.C., Jaeger, K.E., Busch, W., Schmid, M., Lohmann, J.U., and Weigel, D. (2005). Integration of spatial and temporal information during floral induction in Arabidopsis. *Science* 309, 1056-1059.
- Wilczek, A.M., Roe, J.L., Knapp, M.C., Cooper, M.D., Lopez-Gallego, C., Martin, L.J., Muir, C.D., Sim, S., Walker, A., Anderson, J., *et al.* (2009). Effects of Genetic Perturbation on Seasonal Life History Plasticity. *Science* 323, 930-934.
- Willige, B.C., Ghosh, S., Nill, C., Zourelidou, M., Dohmann, E.M., Maier, A., and Schwechheimer, C. (2007). The DELLA domain of GA INSENSITIVE mediates the interaction with the GA INSENSITIVE DWARF1A gibberellin receptor of Arabidopsis. *Plant Cell* 19, 1209-1220.
- Wittkopp, P.J., Haerum, B.K., and Clark, A.G. (2004). Evolutionary changes in cis and trans gene regulation. *Nature* 430, 85-88.
- Wu, G., Park, M.Y., Conway, S.R., Wang, J.W., Weigel, D., and Poethig, R.S. (2009). The sequential action of miR156 and miR172 regulates developmental timing in Arabidopsis. *Cell* 138, 750-759.
- Wu, G., and Poethig, R.S. (2006). Temporal regulation of shoot development in Arabidopsis thaliana by miR156 and its target SPL3. *Development* 133, 3539-3547.
- Yang, L., Xu, M., Koo, Y., He, J., and Poethig, R.S. (2013). Sugar promotes vegetative phase change in Arabidopsis thaliana by repressing the expression of MIR156A and MIR156C. *eLife* 2, e00260.
- Yano, M. (2001). Genetic and molecular dissection of naturally occurring variation. *Curr Opin Plant Biol* 4, 130-135.
- Yoo, S.K., Chung, K.S., Kim, J., Lee, J.H., Hong, S.M., Yoo, S.J., Yoo, S.Y., Lee, J.S., and Ahn, J.H. (2005). CONSTANS activates SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 through FLOWERING LOCUS T to promote flowering in Arabidopsis. *Plant Physiol* 139, 770-778.
- Yu, C.W., Chang, K.Y., and Wu, K. (2016). Genome-Wide Analysis of Gene Regulatory Networks of the FVE-HDA6-FLD Complex in Arabidopsis. *Frontiers in plant science* 7, 555.
- Yu, S., Cao, L., Zhou, C.M., Zhang, T.Q., Lian, H., Sun, Y., Wu, J., Huang, J., Wang, G., and Wang, J.W. (2013). Sugar is an endogenous cue for juvenile-to-adult phase transition in plants. *eLife* 2, e00269.
- Yu, S., Galvao, V.C., Zhang, Y.C., Horrer, D., Zhang, T.Q., Hao, Y.H., Feng, Y.Q., Wang, S., Schmid, M., and Wang, J.W. (2012). Gibberellin regulates the Arabidopsis floral transition through miR156-targeted SQUAMOSA promoter binding-like transcription factors. *Plant Cell* 24, 3320-3332.
- Yu, X., Li, L., Li, L., Guo, M., Chory, J., and Yin, Y. (2008). Modulation of brassinosteroid-regulated gene expression by Jumonji domain-containing proteins ELF6 and REF6 in Arabidopsis. *Proc Natl Acad Sci U S A* 105, 7618-7623.
- Zhai, Q., Zhang, X., Wu, F., Feng, H., Deng, L., Xu, L., Zhang, M., Wang, Q., and Li, C. (2015). Transcriptional Mechanism of Jasmonate Receptor COI1-Mediated Delay of Flowering Time in Arabidopsis. *Plant Cell* 27, 2814-2828.

## Curriculum vitae

### Zhang, Lei

Current address: Tübingerstr 16, Köln, 50739.

Nationality: Singapore

Birth date: 18 May 1983. Birth place: Harbin, China

### Education

- 2014.01 – present Phd International Max Planck Research School (IMPRS), Max Planck Institute for Plant Breeding Research. First two years under supervision of Dr. Jose Jimenez-Gomez (Department of plant breeding genetics); then Prof. Dr. Gero Coupland (Department of plant developmental biology). Cologne, Germany.
- 2011.10 – 2013.12 International Masters Course in Molecular Biology and Biotechnology (CEMBIO), University of Bonn, Germany. Master thesis under the supervision of Dr. Jose Jimenez-Gomez (Max Planck Institute for Plant Breeding Research) and Prof. Dr. Dorothea Bartels (Institute for Molecular Physiology and Biotechnology of Plant, University of Bonn).
- 2003.08 – 2007.06 Bachelor of science in Department of Biological Science, National University of Singapore.
- 2001.10 – 2003.07 Pre-university programmes at National University of Singapore.
- 1999.09 – 2001.09 Harbin Number One Railway Senior Middle School (Harbin, China).

### Work Experience

- 2007.07 – 2009.06 Research assistant in the research group of Prof. Dr. Hao Yu (Department of Biological Science, National University of Singapore).
- 2009.07 – 2010.12 Teaching assistant in the Department of Biological Science, National University of Singapore.

## Acknowledgements

First of all, I would like to express my gratitude to my supervisor Dr. José Jiménez Gómez for offering this research opportunity in his group, and for his patient guidance and helpful comments on my scientific work. Moreover, I really appreciate that he always has an open door and make time for my questions; even after he physically moved to INRA, Versailles. I especially thank him for the encouraging and inspiring conversations when I am feeling low.

Further, I owe my gratitude to Prof. Dr. Maarten Koornneef and Prof. Dr. George Coupland for their interests and valuable advises on my projects. George kindly adopted me to his group and funded my PhD project after the closure of the Koornneef department. Besides, I would like to extend my gratitude to the rest of my thesis committee, Prof. Dr. Ute Höcker, Dr. Maria Albani and Dr. Korbinian Schneeberger.

I would also like to thank our collaborator Dr. Gabriel Schaaf from University of Tübingen on his help in the VIH2 project. Thanks also go to Dr Ivan Acosta, who provide us valuable suggestions and mutant material for the VIH2 project.

Also, I would like to express my greatest appreciation to many of my colleagues from MPIPZ. Special thanks go to my co-worker and friend Dr. Inga Schmalenbach, who initiated both projects in my thesis. I thank all the members in the former Jiménez-Gómez group, Albani group, Pecinka group, Soppe group and Coupland group for their scientific input, but also for a nice working environment and wonderful time outside work.

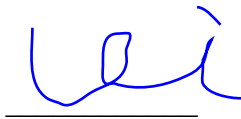
Thanks to the IMPRS and the Max-Planck Society for funding of this project

Finally, I deeply thank my parents for the freedom they gave me to peruse my career in science. A very deep and special gratitude goes to my beloved boyfriend, Joscha Muck, who always cheers me up and supported me tremendously in the end phase of this thesis

## Erklärung

"Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen – , die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie – abgesehen von unten angegebenen Teilpublikationen – noch nicht veröffentlicht worden ist, sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. George Coupland betreut worden."

Köln, 22 May 2017



Teilpublikationen

**Functional analysis of the Landsberg *erecta* allele of *FRIGIDA*.**

Inga Schmalenbach, Lei Zhang, Malgorzata Rynagajlo and José M Jiménez-Gómez (2014)  
BMC Plant Biol. 2014 Aug 13;14:218

Weitere Publikationen

**The relationship between flowering time and growth responses to drought in the Arabidopsis Landsberg *erecta* x Antwerp-1 population**

Inga Schmalenbach, Lei Zhang, Matthieu Reymond and José M. Jiménez-Gómez (2014)  
Front Plant Sci. 2014 Nov 11;5:609.